Available online at www.sciencedirect.com
ScienceDirect
Journal of Financial Economics 85 (2007) 857–887
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/jfec

Order flow, dealer profitability, and price formation ☆

Peter Locke[a,*], Zhan Onayev[b]

[a]*Neeley School of Business, TCU, Fort Worth, TX 76129, USA*
[b]*State Street Global Advisors, Boston, MA 02111, USA*

Received 16 March 2005; received in revised form 5 April 2006; accepted 26 May 2006
Available online 18 April 2007

Abstract

We analyze the dynamics of the S&P 500 futures price, finding both short- and long-run effects of order flow on price. While price moves strongly with the order flow in the short-run, the long-run impact is slightly negative, attributable to costly slippage from a hedging propensity in futures markets. We find strong evidence of a state dependence in the relation between price and order flow, using both volume and floor trader income measures as states. We also find that both the long- and short-run impacts of order flow are greater when dealer income is higher.

JEL classifications: G12; G20

Keywords: Trading; Order flow; Information; Dealers

1. Introduction

We analyze the dynamic link between the S&P 500 futures price and the futures order flow. Our study employs a high-frequency analysis of transactions data to investigate whether the relation between price and order flow varies in different information states of

☆ The paper has benefited from presentations at the George Washington University, Kent State University, the M. J. Neeley School of Business at TCU, and the Financial Management Association. Locke thanks the Crain family for supporting this research while he was at the George Washington University. The views and opinions expressed are those of the authors and do not necessarily reflect those of State Street Global Advisors.

*Corresponding author. Fax: +1 817 921 7227.
E-mail address: plocke@tcu.edu (P. Locke).

0304-405X/$ - see front matter
doi:10.1016/j.jfineco.2006.05.010

the market as represented by volume and floor trader income. The effect of order flow on price has been extensively studied in equity markets (e.g., Lakonishok, Shleifer, and Vishny, 1992; Wermers, 1999; Stoll, 2000; Chordia, Roll, and Subrahmanyam, 2002) and foreign exchange markets (e.g., Evans, 2002; Evans and Lyons, 2002). In both equity and FX studies, order flow conveys otherwise unobservable private information to market participants. The general finding of these studies is that order flow has a permanent effect on price, suggesting a link between information, order flow, and, by extension, dealer profitability.[1]

While there are a large number of studies that document a permanent price impact from order flow in FX and securities markets, there are only a few studies that examine the relation between price and order flow in futures markets. Ferguson, Mann, and Wetsburd (2004) analyze the impact of customer trades on futures prices and find that they have a larger impact on price than other types of trades. These findings are consistent with Manaster and Mann (1996, 1999), Ferguson and Mann (2001), and Kurov and Lasser (2004). On the other hand, Chakravarty and Li (2003) find that futures floor traders, when dual trading, do not profit from their access to order flow. This seems to leave open the question of how (or even what) futures floor traders learn from order flow, and this paper continues that quest.

Order flow generally means changes in dealer inventory, and a companion question to dealer information processing is dealer inventory management. Here, previous results regarding futures floor traders' inventory management are particularly interesting. Manaster and Mann (1996) provide evidence of weak inventory effects in futures markets: trader inventories mean revert, but the relation to price seems counterintuitive in that on average traders unwind positions at favorable prices. Further evidence is provided by Ferguson and Mann (2001), who show that while average futures spreads are similar in magnitude to those in other markets, effective spreads are often negative. The authors interpret these results as evidence that floor traders are often proactive traders, rather than passive order fillers, who might take positions based on inferences from and predictions about order flow, using a more elaborate analysis than that assumed in typical microstructure models. Futures floor traders have an informational advantage over customers due to the ability associated with their physical location to infer the pending order flow of off-exchange customers. Note that if this advantage derives from observing the phone activity on the floor and general hand waving and shouting, it need not be fraudulent, as would be the case if a particular dealer had access to particular pending customer orders. Chakravarty and Li (2003) find no evidence of such particular advantages. Kurov and Lasser (2004) provide further support for Ferguson and Mann (2001). They show that trades initiated by floor traders have a greater impact on price than those initiated by off-exchange customers, strengthening evidence in favor of the existence of trading in these markets related to inferences of anticipated order flow. In the absence of reported quotes, Kurov and Lasser (2004) infer trade initiation by the price of a trade relative to recent prices. Kurov and Lasser (2004) build on Ready (1999), who shows that

[1]In traditional microstructure models (e.g., Kyle, 1985; Glosten and Milgrom, 1985), transactions noisily convey private fundamental information, thus determining the dynamics of asset prices. Studies that analyze order flow in equity markets include Kraus and Stoll (1972), Lee (1992), and Stoll (2000). Chordia, Roll, and Subrahmanyam (2002) examine a cross-section of NYSE stocks for a ten-year period and find a positive relation between lagged order flow and returns.

NYSE specialist discretionary participation is clearly informative. While these studies provide support for some type of information-based floor trading in futures markets, none of them indicates the source of the information, the type of information, or whether the effect of order flow on price is permanent or transient.

The S&P 500 index futures price is linked to the current value of the S&P 500 index through a zero-arbitrage relation. Thus, information (other than changes in the cost of carry, i.e., near-term dividends and the short-term interest rate) related to the S&P 500 index futures is also information about the S&P 500 index. If the S&P 500 futures order flow proves to be a source of information, this can be interpreted as an indication that at least some index trading in general is related to information, which we assume to be private information such as changes in risk preferences, or perhaps very costly macroeconomic research.

The term "order flow" has taken on different meanings in different settings. The common feature of these definitions is a measure of the net trading of some group (e.g., customers) or the inference of trade direction through components of some mechanism trading at quoted prices). For example, some research uses the implied difference between trading at bid and ask quotes to calculate an order flow, irrespective of who the trade participants might be. We are interested in the extent to which futures floor traders provide a price discovery function, and hence the order flow we are interested in is the net proprietary trading by futures customers on the floor. Of course, in addition to the S&P 500 futures there is much more trading in S&P 500 stocks and mutual funds as well as in exchange-traded funds and other derivative instruments based on the S&P 500, including the 100% electronic E-mini S&P 500 futures on the Chicago Mercantile Exchange (CME). The E-mini, one-fifth the size of the floor-traded S&P 500, has gained in importance as shown, for example, by Kurov and Lasser (2004), although this growth has occurred beyond the timeframe of our data.The trading in this plethora of S&P 500 surrogates clearly biases against our finding any permanent price effect from customers trading in S&P futures.

We also seek to learn the extent to which both the permanent impact and the short-run sensitivity of price to order flow vary in different market states, measured by trading volume and dealer profitability. This is motivated by FX studies (e.g., Evans, 2002; Luo, 2001) that find a nonlinear relation between order flow and exchange rate movements. For example, Luo (2001) finds that order flow has a larger impact in periods of large bid-ask spreads, high volatility, and low trading volume. Since the order flow in equity futures markets, similar to the FX market, is non-constant and arrives with a degree of periodicity, we expect to find state-dependency in the relation between order flow and price. We first use the traditional measure of the information flow, trade volume. We next employ a measure of floor trader income, the cash flow to and from floor traders' proprietary accounts, as a proxy for information flows.

The use of volume as an information proxy has been studied extensively. For example, the intraday variation in equity and FX volume has been attributed to a change in liquidity demand or a time-varying percentage of asymmetric information. Recently, Chae (2005) provides evidence that volume decreases before and increases after scheduled corporate announcements. These results are based on the strategic informed trader models of Admati and Pfleiderer (1988) and Foster and Viswanathan (1990), where information *per trade* declines with volume. One of our contributions is to introduce a high-frequency measure of floor trader income as a measure of the information flow. To be consistent with traditional microstructure models, we should find that dealers earn larger profits when the order flow is unusually less informative. For example, the Kyle (1985) dealer forms a supply of liquidity function based on a rationally expected level of information and flow of noise trading. Conditional on this fixed supply function, if liquidity-based trading is unusually high (low), then dealers come out ahead (behind), while on average, dealers break even. In these types of models, the dealer mechanism transfers money from uninformed customers to those better informed.

There are other possibilities regarding the relations between order flow and information. In Easley and O'Hara (1992), due to the assumption of a trickle of information, dealers could earn higher revenue when the order flow is most informative. Similarly, Ferguson and Mann (2001) suggest that futures floor traders might be able to infer semifundamental private information due to their privileged physical access to order flow, and thus their revenue should be higher at times of a more informative order flow. Semifundamental private information, a term developed by Ito, Lyons, and Melvin (1998), refers to knowledge about such things as other market participants' short-term trading strategies and objectives. For example, knowledge about inventories across traders and associated inventory management techniques can be considered semifundamental information. If floor traders actually infer semifundamental information from the market, then this should affect the information and floor trader income relation in a manner different than that implied by traditional microstructure.

To investigate these relations, we use a high-frequency measure of dealer income and decompose this into two components, spread and speculative income, expanding on a technique used in Manaster and Mann (1999) and Chakravarty and Li (2003). Spread income is representative of a dealer's order processing income. Speculative dealer income is the income earned from carrying inventory across time periods. In this latter measure, a temporal element is added to dealing, allowing for a speculative strategy in addition to the traditional, mechanical role of liquidity provision. According to Easley and O'Hara (1992), and consistent with the predictions of Ferguson and Mann (2001), speculative income would vary in different information states, such that dealer income would rise with an increased information flow.[2]

We analyze the interaction between trading volume, dealer profitability, and the informativeness of order flow with the help of a logistic smooth transition regression (LSTR) framework, as in Granger and Teräsvirta (1993), Teräsvirta (1994), and Van Dijk, Teräsvirta, and Franses (2000). This framework allows for two regimes, associated with extreme information states, where the transition from one regime to the other is smooth. On the other hand, the LSTR model can be said to allow for a continuum of regimes. The regimes are associated with changes in the sensitivity of price to order flow, with measures of volume and floor trader income identifying the market states.

Our first result establishes that while the temporary effect of customer buy orders on price is large and positive, the permanent effect is small and negative. The transient effect is as expected: in the presence of an increased order flow floor (customer buy pressure),

[2]We aggregate trading across floor traders, so the speculation we infer by dealer positions is a marketwide or typical dealer phenomenon. Also note that on the futures floor, a five-minute interval (or shorter) will provide enough time to observe some offsetting trades as well as speculation. The higher the frequency, the more imbalanced the trading will be (i.e., more "speculation"), and the richer our analysis. We settled on five minutes after examining a frequency of one minute and finding too many observations with no trading.

trades occur at substantially higher prices. The slightly negative (though significant) long-run effect might seem odd in light of previous literature showing that the S&P 500 futures leads the cash market. Given the lead of futures over cash, one might expect to find informed trading in the futures, and hence a positive long-run relation between order flow and the futures price. However, the arbitrage and quasi-arbitrage motives that link the cash and futures markets and drive the intermarket dynamics are independent of the information structure. In particular, transient as well as permanent price movements will be transferred across markets through arbitrage.

Nevertheless, there remains a slight puzzle regarding our finding of a negative long run relation between order flow and price. One explanation relies on a traditional hedging motive for customer trading, on average, in the S&P 500 futures. The traditional theory of hedging in futures markets, such as the early research by Working (1953), is that derivatives such as these are used as temporary substitutes for cash positions, with the timing of the trades related to some underlying commercial activity. Trades such as anticipatory hedge trades and their offset by pension funds or other indexed funds are considered risk reducing, and thus the hedger is expected to pay for this reduction in the form of some price slippage or long-run contrary movement. While this type of reasoning seems easiest to apply to agricultural markets such as pork bellies and soybeans, we should not be surprised to find such a similar effect in financial futures. In our results, the long-run negative price impact of the customer order flow is consistent with such slippage, or costly trading.

Our second result is that both the short-run and long-run price impacts vary with the flow of information, with volume and floor trader income as our proxies for information. For example, when volume is high (low), order flow has a strong positive (negative) permanent price impact. Our novel finding is that price dynamics also vary significantly with floor trader income, with this finding especially noticeable for the short-run price impact. For example, we show that when floor trader income is high, the temporary impact of order flow on price shifts from the contemporaneous order flow to the lead, suggesting higher profitability associated with an anticipation of the order flow by floor traders. This anticipation of semifundamental information, in the form of customer trading intentions, is suggested by Ferguson and Mann (2001) and Kurov (2005). We also find that the permanent order flow effect is larger and positive at times when floor trader income is highest.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 presents our empirical model and estimation techniques. Section 4 reports results of the estimation. Section 5 concludes.

2. Data

Our analysis uses S&P 500 futures market transactions data over the period from January 1998 through December 2001. Through the generosity of the Commodity Futures Trading Commission we have obtained transactions records from the S&P 500 and E-mini S&P 500 futures contracts. Trades are executed exclusively by members (or their leaseholders). Each transaction record provided to us contains the date and time of the trade, the delivery month, the trade direction (i.e., whether the trader is buying or selling), the classification of the customer type (trade for the member's account, his house's account, another member on the floor, or a customer) and a similar classification for the trade counterparty, the trade quantity, and the price in index points.

Exchange members physically located in the common trading floor have the unique ability to execute proprietary trades on the floor. Other traders wishing to have trades executed on the floor must communicate these orders to agents on the floor to have them executed by members acting as a broker. Proprietary trading by members is limited to the interior of the arena-shaped pit—they must be off of the "top step," and hence not in view of the customer orders flashed in from phone desks. S&P 500 futures floor trading has a daily open and close, with trading hours from 8:30 a.m. to 3:15 p.m. Chicago time, Monday through Friday. At other times, S&P 500 futures trading is relegated to the GLOBEX electronic trading system. The E-mini S&P 500 (one-fifth the size of its big brother) trades exclusively on GLOBEX, and almost 24 hours a day. GLOBEX offers a substantially different market structure, and a different information structure, with no floor traders acting as dealers.[3] In spite of the availability of GLOBEX, which operates 24 hours, seven days a week, market activity is sparse on weekends and holidays, and non-floor trading hours. In other words, S&P 500 index futures trading is most active when US equities are most active, and we focus our empirical analysis on this period.

Our dataset has two valuable characteristics: a long sample period (the four years from January 1998 through December 2001) and important information about the quantity and type of principal behind each side of each trade. Many FX studies (e.g., Luo, 2001) lack information about the size of the trade, forcing the assumption that, regardless of their size, trades have equal impact on prices. On the other hand, because of the nature of futures floor trading, our data lack ex ante quote information, since quotes are only shouted, have an infinitesimally short shelf life, and are not recorded. The only information available to us is ex post transaction price and quantity, with the associated trader information.

3. Basic model and estimation techniques

3.1. Empirical model

Development of the empirical model requires some description of the institutional details of futures trading and a discussion of possible information flows. Floor trading of the S&P 500 futures market is a centralized multiple-dealer market where trades are executed by exchange members in plain view. The members present can execute trades for others, as agents, or for their own accounts, subject to natural priorities and fiduciary responsibilities. Our model is adapted from Evans (2002), who applies it to FX trading, and it might be helpful at this point to also compare FX and futures trading architectures.

While the FX market is mostly decentralized, with dealers physically separated from each other and transactions made by telephone, telex, or computer networks, the S&P 500 futures market is highly centralized. Futures trades are executed on the floor at publicly available prices and all market participants face the same trading opportunities. In addition, FX trading volume is conducted mainly among dealers, usually large banks, with substantial client contacts, including other banks and central banks. In the futures market,

[3] Floor traders can use GLOBEX terminals to enter brokered E-mini trades during regular floor hours, but even these proprietary trades by members on the floor can only occur in the "pit" via open outcry. In particular, for the S&P 500 contract and the E-mini version, these proprietary trades occur down in the pit, off of the "top-step," and thus with a limited view of the flashing of the orders that constitute the bulk of brokered trades.

the floor traders are in close physical contact, isolated on the floor, and have little access to customer information, instead basing their trading strategies on the ebb and flow of orders and prices.

As a consequence of the more centralized trading environment, the S&P 500 futures market might be considered more integrated and transparent than the FX market. In the FX market, not all dealer quotes are observable and transactions can therefore occur simultaneously at different prices, although because of the sequential nature of search among these prices, interdealer arbitrage is unlikely. This argument is supported by Evans (2002) who shows that the lack of transparency associated with direct FX trading leads to the existence of an equilibrium distribution of transaction prices across dealers at any point of time. It is this distribution of prices that is modeled by Evans as a source of transient noise in FX prices. In the futures market, while the latest transaction price is publicly known through the ‘‘time and sales’’ collection and broadcast process, there remains the possibility that transactions can occur on the floor at different prices at almost the same point in time. In fact, the trading floor is often described as chaotic, and the difference between simultaneous transactions prices may reflect search cost associated with the floor trading, similar to the FX distribution across of prices across multiple dealers. Unlike the FX market, futures dealers are located on one trading floor, and can observe other dealers' behavior and make some assessment regarding other market participants' positions. However, these are likely to be very noisy assessments. Overall, there is still plenty of room for a good deal of uncertainty, in equilibrium, among the floor traders.

The information structure of the S&P 500 futures market resembles that of the FX market. Both public and private information have potential effects on price formation in all financial markets, albeit with different magnitudes across markets. For instance, in equity markets, private fundamental news, i.e., information specific to a company's cash flows, is assumed to play an important role in price formation, while in the FX or government bond markets there is typically no assumption of such fundamental information that might affect a dealer's pricing behavior (Cao, Evans, and Lyons, 2006). A similar argument holds for financial futures.

Working with the assumption of an equilibrium, arbitrage-free distribution of prices across floor traders, we proceed to describe the S&P 500 index futures information environment. First, fundamental public news symmetrically shifts values, and hence the distribution of all S&P 500 futures transaction prices. Examples of this kind of public news are macroeconomic announcements or price movements in related securities such as government bonds, oil futures, etc. Second, the distribution of non-fundamental public information can be either symmetric or asymmetric. This distinction depends on whether all floor traders and customers interpret public signals homogeneously or heterogeneously. If a significant divergence of opinion occurs, then the price distribution pattern can be altered. Third, private fundamental information, i.e., superior non-public information about a security's future cash flows, probably does not play a very important role in price formation. Because the S&P 500 index consists of multiple securities with uncertain cash flows, it is highly improbable for someone to be able predict, especially at a high frequency, innovations to the index's future cash flows based on private fundamental information. It would certainly be costly to obtain sufficient valuable private information at a high frequency on the 500 component stocks. This feature of the information environment of the futures market is analogous to that of FX markets. Finally, private non-fundamental information, superior information unrelated to expected cash flows (e.g., change in risk presences, endowments, trading constraints, etc.), plays an important role in the price formation in the futures markets. Floor traders' proximity to the order flow could provide them with superior information that is orthogonal to fundamentals (see Cao, Evnas, and Lyons, 2006). Since every floor trader observes only a fraction of the order flow, the interpretation of the signal that follows differs across the floor. This information component, mainly private information that is conveyed noisily by the customer order flow, is expected to contribute most to changes in the pattern of transaction prices.

Based on this framework, we follow Evans (2002) and assume that information arrives in two forms: common knowledge news and non-common knowledge news. Common knowledge news is characterized by the simultaneous arrival of information to all market participants, floor traders, and customers, and a homogeneous interpretation across these participants. This news could be in the form of public announcements that might affect the S&P 500 but are not likely to cause much of a divergence of opinion among traders. Non-common knowledge news can be generated by both public and private sources. Public news announcements that lead to different interpretations across traders fall into this category. Finally, private non-common knowledge news is mainly non-fundamental or semifundamental information. In summary, non-common knowledge news affects the size or direction of trading and possibly both, providing for a systematic and permanent relation between order flow and price changes. On the other hand, the common knowledge news component causes an immediate effect on all transaction prices, since all traders update their beliefs about the fair value of the asset simultaneously, and thus it is expected to have no relation to order flow. The relative effect of these two sources of price innovation depends on the market structure as well as the extent of disagreement on the futures floor.

From trading on the S&P 500 floor, transaction records are generated with associated prices and quantities, as described above. We represent the logarithm of the price of one of these transactions, involving a floor trader and a customer, observed at time t, p_t^o, by

$$p_t^o = p_t + \omega_t^o \tag{1}$$

for $o = \{a, b\}$. The values ω_t^a and ω_t^b represent idiosyncratic shocks by which observed floor trader personal sale $\left(\omega_t^a\right)$ and purchase $\left(\omega_t^b\right)$ prices (in logs) differ from the latent common market value, which we denote p_t. The innovations ω_t^a and ω_t^b are serially uncorrelated and independently distributed. We use the superscripts a and b to correspond to the traditional microstructure literature where dealers set quotes to trade at *asks* or offers and *bids*. In the S&P 500 futures double oral auction, with floor traders simultaneously acting as brokers and dealers trading at many different prices, and with no reported quotes, we use the realized trading by customers, rather than trading at quotes, to correspond to trading at an ask or a bid. Thus, in our setup the superscript a indicates a transaction involving the purchase by a customer, and a b indicates a transaction involving the sale by a customer.

Both common and non-common knowledge shocks lead to a change in the equilibrium price. The relation between the latent equilibrium returns, i.e., the change in p_t, and the components—common knowledge news, which we denote ε_t, and non-common news, which we denote v_t—is described by

$$\Delta p_t = B(L)\, v_t + \varepsilon_t, \tag{2}$$

$$x_t = C(L) v_t, \tag{3}$$

where x_t is the order flow, with $x_t > 0$ indicating net customer buying. The functions $B(L)$ and $C(L)$ are polynomials in the lag operator, L (these functions can include leads and lags). The innovations, ε_t and v_t, are mutually independent and serially uncorrelated, and are independent from the idiosyncratic shocks $\left(\omega_t^a \text{ and } \omega_t^b\right)$. Combining the relations from (1) and (2) we can describe the observed return, or the change in the log of the observed price over an interval, by

$$\mu_t^o = \Delta p_t^o = \varepsilon_t + B(L)v_t + \omega_t^o - \omega_{t-1}^o, \tag{4}$$

for $o = \{a, b\}$. Finally, the bivariate relation between observed returns and order flow is

$$\begin{bmatrix} \mu_t^a \\ \mu_t^b \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} D(L)x_t + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \varepsilon_t + \begin{bmatrix} \omega_t^a - \omega_{t-1}^a \\ \omega_t^b - \omega_{t-1}^b \end{bmatrix}, \tag{5}$$

where $D(L) = B(L)C(L)^{-1}$. We allow for a general specification of the $D(L)$ polynomial. The sum of the coefficients of the polynomial $D(L)$, D^*, yields the long-run impact of order flow on price. The polynomials $B(L)$ and $C(L)$ can take many forms, provided that $C(L)$ is invertible, with these forms describing the dynamic responses of prices and order flow to non-common knowledge shocks.

In Eq. (5), the order flow, x_t, is driven entirely by non-common knowledge news, the v_t that has been eliminated by substitution. On the other hand, common knowledge shocks, represented by ε_t, only shift the distribution of price changes, and do not impact the order flow. In the presence of a common knowledge shock, dealers will seek to avoid the risk associated with large inventory imbalances by moving both buy and sell quotes in tandem. For example, if there is public information suggesting that the S&P 500 should be above the current quoted prices, floor traders will increase both their bids and offers. Customers will receive exactly the same public information and, faced with the shift in quotes, will not change their net trading pattern. Overall, while common knowledge news can impact the level of volume, perhaps due to hedging or portfolio rebalancing, the sign of the order flow should not be predicted by changes in common knowledge.

3.2. Calculation of returns and order flow

We divide each trading day into five-minute intervals. We obtain the last transaction customer-to-dealer purchase and sale prices (in logarithms) recorded during each interval t, p_t^a and p_t^b, respectively, are compute the differences in their natural logarithms to obtain the returns series $\mu_t^a = p_t^a - p_{t-1}^a$ and $\mu_t^b = p_t^b - p_{t-1}^b$. On any day, only the most active futures contracts are used for obtaining the price series. The order flow, x_t, is measured for each five-minute interval as net customer buying. Several volume and floor trader income measures, described below, are also computed for each five-minute interval t. Note that with futures floor trading, all orders are processed on the floor, and must be executed by a member. Customer orders are either flashed via hand signals into the pit for immediate execution or held as "paper," price-contingent orders, on the floor by a member. Proprietary trading by members is entirely discretionary, with no obligation to provide an orderly market, a narrow spread, or any requirement to quote at all or even be present on the floor let alone in any particular futures pit.

As stated above, we limit the analysis to floor trading hours and designate all overnight observations as missing. To do otherwise would confound the estimated high-frequency covariances. Therefore, when computing any value of a variable, if a particular variable involves a value (possibly its own lagged value) that is designated as missing, then we also designate the observation as missing.

3.3. Dealer income decomposition

We next describe our calculation of the high-frequency measure of dealer income and its decomposition. The methodology we use provides an unbiased high-frequency measure of aggregate floor trader income and decomposes it into spread income and speculative income. Spread income is the income earned when floor traders offset a balanced set of trades within a short interval, comparable to the results of trading at a bid-ask spread. Speculative income comes from longer-term trading, perhaps based on information from incoming orders. Speculative income arises only when floor traders take a strategic position, rather than offsetting positions within an interval.

In each five-minute interval, floor traders execute trades for their personal accounts against the order flow from off-exchange customers. We measure the income earned by all floor traders in an interval by marking each trade to market using the end-of-day settlement price, as in Chakravarty and Li (2003). Manaster and Mann (1999) use this method to calculate daily income and similarly decompose it into daily spread and speculative income.

Consider the following example. During interval t, dealers sell five contracts at \$80, and buy four contracts at \$70 for their personal accounts. Without further assumptions, the definition of dealer income in this situation is unclear. One interpretation is that they have earned \$40 and have an inventory of one contract that has not led to income. But our concern is the interpretation of these end-of-interval inventories in terms of their information flow. Suppose in the subsequent interval, $t+1$, dealers buy three contracts for \$75 and sell two contracts for \$75. Again, without further assumptions, income in this interval is unclear. Using the same accounting method, in this market the dealer would have an income of zero and an inventory change of -1 contract. We account directly for the imbalances by using the settlement price, the price eventually used by the exchange to calculate daily cash flows for dealers (Chakravarty and Li, 2003). Continuing with the example, suppose that the settlement price on this day is \$75. Then, using marking-to-settlement, for all trades, total income in interval t is \$45, or \$9 per contract, and the income in interval $t+1$ is zero. Also in t, spread income is \$40, or \$8 per contract, and speculative income is \$5, or \$1 per contract.

We compute the income, I_t, for each five-minute interval, t, of each trading day (omitting for convenience in this notation a reference to the particular day). The marking to market leads to the following income calculation:

$$I_t = \sum_{i=1}^{N_t} Q_{Bti}(K - P_{Bti}) + \sum_{j=1}^{M_t} Q_{Stj}(P_{Stj} - K), \tag{6}$$

where Q_{Bti} is the buy quantity for a floor trader for transaction i in interval t, P_{Bti} is the buy price for a floor trader for transaction i in interval t, and $i = 1,\ldots,N_t$ is an index of floor trader buy trades in interval t. Similarly, Q_{Stj} is the sell quantity for a floor trader for transaction j in interval t, P_{Stj} is the sell price for a floor trader for transaction j in interval t, and $j = 1,\ldots,M_t$ is an index of floor trader sell trades in interval t. K is the daily settlement price.

We decompose the income of that interval into two components, spread income and speculative income. First define $NET_t = \sum_{i=1}^{N_t} Q_{Bti} - \sum_{j=1}^{M_t} Q_{Stj}$. This gives a measure of trades which have not been offset within an interval, with $NET_t > 0$ indicating that floor traders have bought more contracts than they have sold. Spread income, S_t, is given by

$$S_t = \sum_{i=1}^{N_t} Q_{Bti}\left(\overline{P}_{St} - \overline{P}_{Bt}\right) \quad \text{if} \quad NET_t < 0, \tag{7}$$

$$\text{and} \quad S_t = \sum_{i=1}^{N_t} Q_{Sti}\left(\overline{P}_{St} - \overline{P}_{Bt}\right) \quad \text{if} \quad NET_t > 0,$$

where $\overline{P}_{St}$ and $\overline{P}_{Bt}$ are the quantity-weighted average sell and purchase price for floor traders in the interval. Similarly, speculative income in the interval, Y_t, is defined by

$$Y_t = \left(\overline{P}_{St} - K\right) * NET_t \quad \text{if} \quad NET_t < 0, \text{ or,} \quad Y_t = \left(K - \overline{P}_{Bt}\right) * NET_t \quad \text{if} \quad NET_t > 0. \tag{8}$$

We normalize these component income measures, dividing by the total number of contracts offset, which is the maximum of either the quantity sold or quantity bought in the interval. Thus, in this high-frequency application of the technique used in Manaster and Mann (1999), total income per contract is allocated among two components, a spread component due to offsetting trading within an interval, and a speculative component due to holding a position for a longer time.

In this framework, both speculative and spread income per contract (and hence total dealer income) are unconstrained, and depend on the dealers' speculative abilities as well as order processing ability. Microstructure models such as Kyle (1985) and Glosten and Milgrom (1985) offer sources of spread income, which typically show it depending on volatility and information asymmetries. Speculative income per contract is due to a deviation of price from the implied equilibrium price. In traditional models this is assumed to be zero, on average, if floor traders have an unbiased expectation of the intrinsic value, K. Manaster and Mann (1999) find that speculative income is the major source of futures dealer income. If floor trader speculation varies with the information flow, variations in speculative income should be indicative of variations in the information flow.

4. Results

4.1. Analysis of order flow and returns

In this section we describe the data and present initial results on the relationship of order flow to price changes. First we discuss the trading day. While the advent of electronic trading has made it possible to trade 24 hours a day, the vast majority of S&P 500 futures transactions in our sample take place during the time when US equity trading occurs. This gives rise to extreme 24-hour intraday patterns in volume and price volatility, and motivates the choice of estimation technique. Evans (2002) shows that the foreign exchange market, which also operates 24 hours a day, exhibits strong intraday patterns in trade intensity, defined as the number of transactions per minute. We create a similar measure of trade intensity, the number of trades per five-minute time interval, and find the expected patterns. Fig. 1 reports the average trade intensity for each five-minute interval for the period between January 1998 and December 2001. Similar to Evans (2002), we find a three-humped pattern in trade intensity. The first small hump occurs around 8 a.m. Chicago time, around the time fixed income trading commences in the US The next hump occurs from 8:30 to 9:00 a.m. At 9:00 a.m. the concentration of trading is the highest. This time corresponds to the first half-hour of regular US equity trading. Finally, the last hump is between 3:00 and 4:00 pm. This time is associated with cessation of regular US equities trading. We also examine trade intensity for different subsamples of our dataset. The smallest sub-sample we examine is three months. While unreported due to obvious space concerns, the patterns remain highly persistent.

Fig. 2 reports the concentration of total trading volume that is calculated as the sum of all floor-traded contracts and normalized E-mini contracts during each five-minute interval. (The quantity of an E-mini trade is normalized by multiplying by 0.2, since it is one fifth the size of the regular S&P.) On average, during regular trading hours, trades are larger than during the overnight period, so that volume has an even more pronounced intraday pattern than trade intensity. From the analysis above it follows that trading activity is highly concentrated during floor trading hours. During hours that the floor is closed, many fewer trades are executed. We exclude the non-floor trading period from our analysis, focusing on the floor trading period, from 8:30 a.m. to 3:15 p.m. Chicago time.

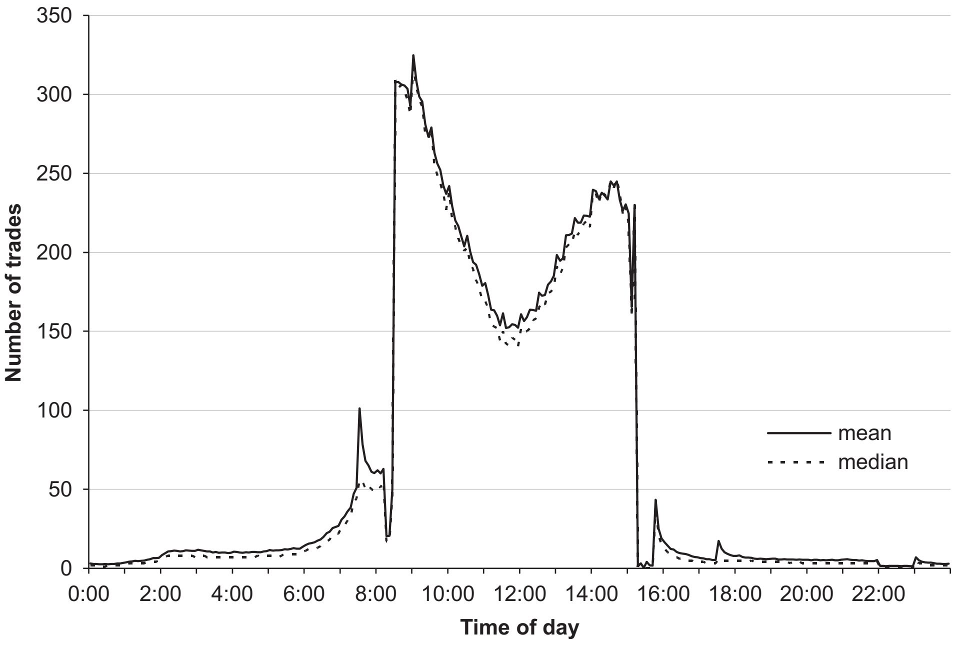

Fig. 1. Trade intensity. The data are for S&P 500 futures trading from January 1998 through December 2001. Trade intensity is the number of transactions per five-minute time interval. Mean and median number of trades are computed by each five-minute time interval across the whole sample period.

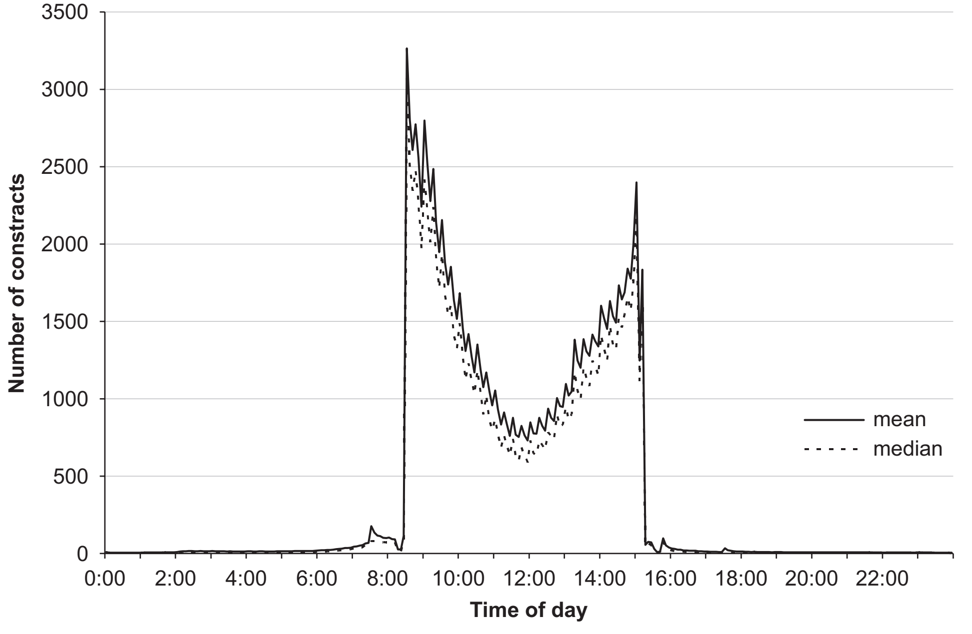

Fig. 2. Trading Volume. The data are S&P 500 futures trading from January 1998 through December 2001. The number of contracts is the volume of trades executed per five-minute time interval. Mean and median volume are computed by each interval across the entire sample period.

Panel A of Table 1 reports some sample statistics. The distribution of the variables informs our choice of empirical methodology away from simple models. It should be clear that the unconditional distributions of returns and order flow appear non-normal. Both distributions are skewed to the left. The distribution of total volume, which is non-negative, shows the typically much larger mean than median. Similarly, distributions of volumes generated by off-exchange customers and hedgers are right-skewed. Hedge trades are trades for members of the exchange on the floor, such as option market makers using brokered futures to hedge.

In addition, the table presents statistics on our floor trader income measures. The distribution of spread income is highly skewed to the left and is extremely fat-tailed. On the other hand, the distribution of speculative income is somewhat skewed to the right. On average, dealers as a group, for each five-minute time interval, earn \$43.72 per contract. Of this amount \$32.44 comes from speculation, or being long or short longer than five-minutes, similar to the findings of Manaster and Mann (1999). However, note also that median speculative income is −\$0.79, indicating that speculation is indeed risky, with losses in more than half of the intervals. On the other hand, spread income is small and positive for the majority of the intervals, which is consistent with spread income being a rough measure of customer execution costs in the S&P 500 futures market, as used by several studies. Total customer costs could include a loss due to slippage associated with inelastic hedging needs, and these will be in addition to immediate execution costs. In view of the large variation in income shown in Panel A of Table 1, we examine in more detail the behavior of income in the tails of the distribution. Our concern is whether intervals with

Table 1
Sample statistics

Panel A. Sample Distributions

	Mean	Median	Std. dev.	Maximum	Minimum	Skewness	Kurtosis
μ_t^a return	0.000	0.000	0.001	0.037	−0.049	−0.57	60.32
x_t, order flow	−69	−45	116	931	−1,595	−1.82	7.92
Total volume	1,094	898	771	10,186	3	1.86	5.74
Customer volume	766	628	536	8,621	1	1.84	5.89
Hedger volume	199	144	188	2,332	1	2.50	9.54
Total income	43.72	8.74	1,458.63	35,992.66	−32,592.30	0.75	38.38
Spread income	11.13	8.17	45.90	2,566.80	−2,169.51	−3.02	510.02
Speculative income	32.44	−0.79	1,459.59	35,953.38	−32,603.54	0.74	38.30

Panel B. Autocorrelations

	1	2	3	4	5	6	7	8	12	16	20
μ_t^a	−0.110	0.005	0.003	−0.002	0.004	0.003	0.003	0.001	0.006	0.000	−0.003
	0.000	(0.190)	(0.293)	(0.349)	(0.248)	(0.260)	(0.292)	(0.412)	(0.093)	(0.472)	(0.255)
μ_t^b	−0.080	−0.004	−0.003	0.004	0.006	0.000	0.000	0.009	−0.001	0.000	0.001
	0.000	(0.231)	(0.250)	(0.212)	(0.122)	(0.480)	(0.469)	(0.032)	(0.381)	(0.460)	(0.425)
$\lvert\mu_t^a\rvert$	0.273	0.208	0.195	0.192	0.195	0.186	0.188	0.180	0.171	0.163	0.149
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

x_t	0.116	0.125	0.150	0.122	0.113	0.135	0.108	0.098	0.090	0.055	0.034
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Volume	0.657	0.557	0.567	0.502	0.469	0.480	0.432	0.404	0.352	0.241	0.164
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Income	0.092	0.059	0.037	0.041	0.032	0.038	0.013	0.019	0.019	0.021	0.017
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Spread income	0.092	0.051	0.052	0.047	0.051	0.060	0.046	0.016	0.012	0.005	0.026
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)
Spec. income	0.095	0.061	0.039	0.042	0.032	0.039	0.013	0.019	0.019	0.022	0.017
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

The order flow, x_t, is net customer buying during five-minute interval t. The values μ_t^a, μ_t^b are logarithmic changes in the last prices at which dealers sold and bought futures, respectively, within interval t. Volume is the combined trading of all futures trader types. Income is floor trader income calculated by marking all trades in an interval to market at the daily settlement price. Speculative income is calculated using the trade imbalance within an interval and marking it to market at the daily settlement price. Spread income is the income from offsetting trades within an interval. All income measures are per contract. In Panel B, the autocorrelations are computed by GMM. The p-values reported in parentheses are calculated from Wald tests of the null hypothesis of a zero correlation allowing for conditional heteroskedasticity.

large positive and negative incomes are clustered around a particular time period or event rather than spread across the sample. If so, we would need to consider modeling these relations directly. In unreported results, we do not find any time-dependent pattern in the behavior of extreme positive and negative values of income (details available upon request).

The time-series properties of the data are examined in Panel B of Table 1. The autocorrelation coefficients are calculated using a two-step generalized method of moments (GMM). Both μ_t^a and μ_t^b display a negative and statistically significant (at the 1% level) autocorrelation at lag one. After lag one, no autocorrelations are statistically significant. Consistent with the futures market being more transparent and perhaps more efficient than the FX market, we find that the return autocorrelation coefficient is substantially lower at 10% than the 30% reported in Evans (2002). Order flow autocorrelation is positive and significant at the first lag, and then exhibits a slowly decaying pattern. While the returns display a relatively low first-order autocorrelation, the absolute value of returns exhibit large autocorrelations that decay very slowly, i.e., volatility is highly persistent. Volume is similarly persistent.

Based on the autocorrelations presented in Panel B of Table 1, unconditional returns seem to follow an MA(1) process. While the autocorrelations of order flow decay, they remain statistically significant at lag 20. This finding is at odds with the FX market, where order flow is autocorrelated only at lag 1. This difference probably arises because our order flow measure is not inferred based on trade initiation but is based instead on actual buys and sells between dealers and off-exchange customers. This is why, at times of increased trading pressure created by off-exchange customers, we are likely to observe a positive autocorrelation in order flow, which in turn reflects the acquisition and disposition of floor trader inventories. The inventory level will increase until customer pressure is absorbed and prices adjust. Inventory effects are perhaps impossible to detect in the available FX data. The autocorrelations of income measures are positive and statistically significant up to 20 lags. Similar to the order flow results, autocorrelations of all three income metrics are low in magnitude and exhibit a pattern of slow decay.

Next we look at the relation between returns and polynomials of order flow. This allows us to specify a reasonable time-series specification for the bivariate model. In Table 2 we present estimates of various specifications for the polynomial $D(L)$ using μ_t^a, the log change in customer buy prices, as the dependent variable. Unreported results for μ_t^b, the log change in customer sell prices, are very similar. The GMM estimates are computed using a subset of moment conditions and the Newey-West weighting matrix, allowing for heteroskedasticity and an MA(1) process in the errors. As a diagnostic check, we report the l-test statistic (Cumby and Huizinga, 1992) for the null hypothesis of the presence of an MA(1) process in the residuals. From the results, it should be clear that the most significant impact of order flow on returns is through the leading terms of the polynomial. Comparison of 14 different specifications shows that leading-term coefficients are significant up to a lead of 12. When we increase the order of the polynomial to 12 leads, the x_{t-2} term becomes insignificant at the 5% level. The estimates suggest that $D(L)$ is well characterized by Model 11, the polynomial that includes terms from $L^{-11}, L^{-10}, \ldots, L$. All coefficients in this specification are statistically significant at a 5% level and the R^2 statistic is 5.8%. The low R^2 for Model 14, which includes only past order flow, suggests that most of the variation in returns is explained by the leading terms of order flow.

Focusing on Model 11, we see that the coefficient on x_t, describing the contemporaneous relation between order flow and changes in the customer purchase price, is positive and

Table 2
Price change decomposition regressions

		$D(L)$ polynomial coefficients (x100,000)																	Diagnostics		
	int.	x_{t-4}	x_{t-2}	x_{t-2}	x_{t-1}	x_t	x_{t+1}	x_{t+2}	x_{t+3}	x_{t+4}	x_{t+5}	x_{t+6}	x_{t+7}	x_{t+8}	x_{t+9}	x_{t+10}	x_{t+11}	x_{t+12}	R-sq.	SEE	l-test
1	17.939				−0.058	0.101	0.226												0.043	0.001	41.211
	(0.000)				(0.000)	(0.000)	(0.000)														(0.000)
2	13.849			−0.032	−0.049	0.108	0.239	−0.053											0.045	0.001	21.749
	(0.000)			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)													(0.040)
3	10.520		−0.013	−0.026	−0.045	0.119	0.250	−0.051	−0.071										0.049	0.001	16.025
	(0.000)		(0.018)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)												(0.190)
4	8.357	−0.010	−0.010	−0.021	−0.037	0.126	0.259	−0.047	−0.070	−0.057									0.049	0.001	10.970
	(0.000)	(0.036)	(0.059)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)											(0.532)
5	6.779			−0.020	−0.031	0.130	0.259	−0.048	−0.072	−0.060	−0.057								0.054	0.001	12.148
	(0.000)			(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)										(0.434)
6	5.379			−0.018	−0.029	0.133	0.262	−0.044	−0.070	−0.057	−0.056	−0.042							0.054	0.001	14.396
	(0.000)			(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)									0.276
7	4.096			−0.016	−0.028	0.129	0.265	−0.040	−0.068	−0.053	−0.054	−0.041	−0.039						0.055	0.001	7.256
	(0.000)			(0.008)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)								(0.840)
8	3.411			−0.015	−0.027	0.066	0.264	−0.038	−0.067	−0.051	−0.056	−0.040	−0.039	−0.024					0.051	0.001	13.544
	(0.000)			(0.014)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)							(0.331)
9	−0.154			−0.012	−0.030	0.123	0.256	−0.044	−0.069	−0.053	−0.057	−0.042	−0.038	−0.026	−0.023				0.055	0.001	8.278
	(0.434)			(0.032)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)						(0.763)
10	2.489			−0.010	−0.027	0.134	0.264	−0.035	−0.066	−0.047	−0.050	−0.035	−0.040	−0.024	−0.022	−0.020			0.057	0.001	13.510
	(0.003)			(0.043)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)					(0.333)

Table 2 (*continued*)

		D(L) polynomial coefficients (x100,000)																	Diagnostics		
	int.	x_{t-4}	x_{t-2}	x_{t-2}	x_{t-1}	x_t	x_{t+1}	x_{t+2}	x_{t+3}	x_{t+4}	x_{t+5}	x_{t+6}	x_{t+7}	x_{t+8}	x_{t+9}	x_{t+10}	x_{t+11}	x_{t+12}	R-sq.	SEE	l-test
11	2.051			−0.010	−0.025	0.133	0.265	−0.035	−0.067	−0.048	−0.047	−0.034	−0.039	−0.026	−0.021	−0.020	−0.015		0.058	0.001	9.547
	(0.016)			(0.043)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.010)				(0.656)
12	−0.020				−0.025	0.139	0.264	−0.040	−0.070	−0.057	−0.050	−0.048	−0.046	−0.037	−0.028	−0.026	−0.015		0.059	0.001	*15.535*
	(0.491)				(0.012)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				*(0.214)*
13	0.526			−0.009	−0.024	0.127	0.269	−0.037	−0.068	−0.049	−0.048	−0.040	−0.037	−0.025	−0.024	−0.019	−0.018	−0.011	0.061	0.001	*13.704*
	(0.280)			(0.056)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)	(0.042)			*(0.320)*
14	(5.456)			−0.009	−0.033	0.125													0.010	0.001	*8.764*
	(0.125)			(0.079)	(0.000)	(0.000)															*(0.723)*

The table entries are GMM estimates (with p-values below in parentheses) of the coefficients of polynomial $D(L)$, for the regression

$$\mu_t^a = \varepsilon_t + D(L)x_t + \omega_t^a - \omega_{t-1}^a.$$

Order flow, x_t, is net customer buying in interval t. μ_t^a is the logarithmic change in the last customer purchase price from interval *t-1* to t. The sample includes all transactions for the period from January 1998 through December 2001. The GMM estimates and standard errors allow for the presence of conditional heteroskedasticity and an MA(1) error structure. The right-hand column reports the l-test statistic (Cumby and Huizinga, 1992) for the null hypothesis of the presence of an MA(1) process in the residuals.

statistically significant at the 1% level. In other words, consistent with traditional microstructure, in the short run price rises (falls) at times when customers are net buyers (sellers). As in Evans (2002), the coefficient on the lead order flow, x_{t+1}, is also positive and much larger than the contemporaneous coefficient. All higher-order lead coefficients, capturing the relation of returns and future order flow, are negative, so that over time price is moving against the customer order flow (e.g., falling prices are associated with net customer buying). This is consistent with inventory reversion on the floor combined with S&P 500 customer trades being generally uninformed hedger trades, on average. In fact, there appears to be slippage, on average, so that customers incur a cost for their use of S&P futures as hedging vehicles. The negative coefficient on x_{t-1} implies that, on average, positive returns are preceded by floor trader buying. This could be due to a prescient ability for floor traders along the lines suggested by Ferguson and Mann (2001).

We proceed with estimation of the complete bivariate model of Eq. (5). This allows a more careful analysis of the origins of price innovations and whether idiosyncratic shocks or common knowledge news cause a larger variation in returns. The observed floor trader purchase and sale prices are drawn independently from their respective distributions. The idiosyncratic shocks, ω_t^a and ω_t^b, associated with dealer prices are serially uncorrelated and independently distributed. They also are orthogonal to common knowledge news shocks, and independent from leads and lags of order flow.

For this specification we allow $D(L)$ to be the 13th-order polynomial containing terms in $L^{-11}$to L, that is, mostly leading order flow terms, following the results of Model 11 from Table 2. Transaction prices, whether customer buys or sells, have the same variance, Σ_ω. The variance of the common knowledge shocks is Σ_ε. To estimate the structural model, we transform it to its state-space form and obtain GMM estimates using as instruments a vector of order flow variables $z_t + [x_{t+11},\dots,x_{t-1}]$ and moments derived from the covariances of observed returns. We follow the procedure described in Evans (2002), where more detail is available.

Table 3 reports estimates for the bivariate model. The order flow coefficients are similar to those reported in Table 2, revealing a positive short-run effect of order flow on price. In addition, in Table 3 we quantify the permanent effect of order flow on price as D^*, the sum of the coefficients of the polynomial $D(L)$. This permanent effect is negative, i.e., swinging from the large positive contemporaneous coefficient to a small negative permanent effect. When prices are falling, customers are buying, on average, in the long run. One interpretation of this finding relates to slippage. In other words, the S&P 500 futures is a derivative market, and the motives for trading this contract are undoubtedly many. The fact that we pick up, on average, what appears to be costly hedging, with prices moving against the typical customer trade, should not be surprising.

Another interesting finding in Table 3 is that the standard deviation of common knowledge shocks, $\Sigma_\varepsilon^{1/2}$, is about six times larger than that of the sampling (idiosyncratic) shocks, $\Sigma_\omega^{1/2}$, so that it is the common knowledge variability that drives the variation in returns. This is somewhat different than the FX market, where more than 50% (see Evans, 2002) of exchange rate volatility is due to idiosyncratic shocks, or price variation among dealers. The institutional differences clearly lead to differences in the sources of price volatility. Nevertheless, the results our estimation of the bivariate model support the existence of an equilibrium distribution of transaction prices at any point in time across the floor traders in the futures market. Based on this fit, we proceed in the next section with testing the sensitivity of the parameters to changes in the flow of information.

Table 3
Bivariate model, price change, and order flow

					$D(L)$ coefficients (x100,000)							
x_{t-1}	x_t	x_{t+1}	x_{t+2}	x_{t+3}	x_{t+4}	x_{t+5}	x_{t+6}	x_{t+7}	x_{t+8}	x_{t+9}	x_{t+10}	x_{t+11}
−0.027	0.125	0.267	−0.043	−0.073	−0.050	−0.050	−0.040	−0.039	−0.026	−0.025	−0.020	−0.016
(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.010)
$\Sigma_{\omega}^{1/2}=$	21.961		$\Sigma_{\varepsilon}^{1/2}=$	121.509		$D^{*}=$	−0.017			J-stat.	18.725	
	(0.000)			(0.000)			(0.000)					

The table reports GMM estimates of the bivariate model $\begin{bmatrix}\mu_t^a\\ \mu_t^b\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}D(L)x_t+\begin{bmatrix}1\\1\end{bmatrix}\varepsilon_t+\begin{bmatrix}\omega_t^a-\omega_{t-1}^a\\ \omega_t^b-\omega_{t-1}^b\end{bmatrix}$ with μ_t^a and μ_t^b the logarithmic hanges in customer purchase and sale prices over a five-minute interval, and the order flow, x_t, net customer buying in the interval. The polynomial lag operator $D(L)$ contains both leads and lags. Associated p-values are reported in parentheses. The sample includes all transactions for the period of January 1998 through December 2001. The number of non-missing observations is 63,722. The moment conditions include lead and lag order flow $[x_{t-1},...x_{t+11}]$ and the variances of μ_t^a and μ_t^b. The long-run impact of order flow, D^*, is the sum of the coefficients of the polynomial $D(L)$. Asymptotic standard errors are corrected for heteroskedasticity and serial correlation. The also table reports Hansen's J-statistic with its associated p-value.

4.2. Nonlinear models

In this section, we investigate whether the relation between price and order flow is market-state dependent. The results of Tables 2 and 3 show that the futures order flow impacts the price formation process in the short and long run. Now we allow the dynamics of the floor traders' response to order flow to vary across states of the market. We test whether order flow has an equal impact regardless of the level of volume as well as contemporaneous measures of floor trader income. Prior to testing these hypotheses, we present graphs of measures of intraday floor trader profitability, and offer some descriptive statistics regarding returns and order flow.

Figs. 3–5 present the intraday behavior of the median per-contract total income, speculative income, and spread income, respectively. From the graphs, it is evident that morning appears to be the most profitable time period for floor traders. At the opening, floor traders as a group earn median income of \$50 per contract. More than half of this income, about \$32, is attributed to profits from speculation, with the remaining \$18 being spread income. This pattern could arise if informed and uninformed noise traders combine their trades in the morning, as suggested in Admati and Pfleiderer (1988). Not surprisingly, most of the end-of-the-day income comes from the spread component. In addition, the pattern shows that as the trading day progresses, the level of speculative income falls, while spread income remains relatively stable. As shown in Fig. 5, spread income is U-shaped and is much less volatile than income from speculation, as we should expect. The U-shaped form suggests that dealers quickly offset morning trades on average at favorable prices, deriving higher positive spread income. At the end of the day, increased liquidity trading by customers could cause the observed jump in spread income. Since the time interval chosen for our analysis is very short, five minutes, we have speculation as a large component of aggregate income. A higher frequency would exacerbate this, clearly, with analysis at the trade-by-trade level causing each trade to be seen as a speculative

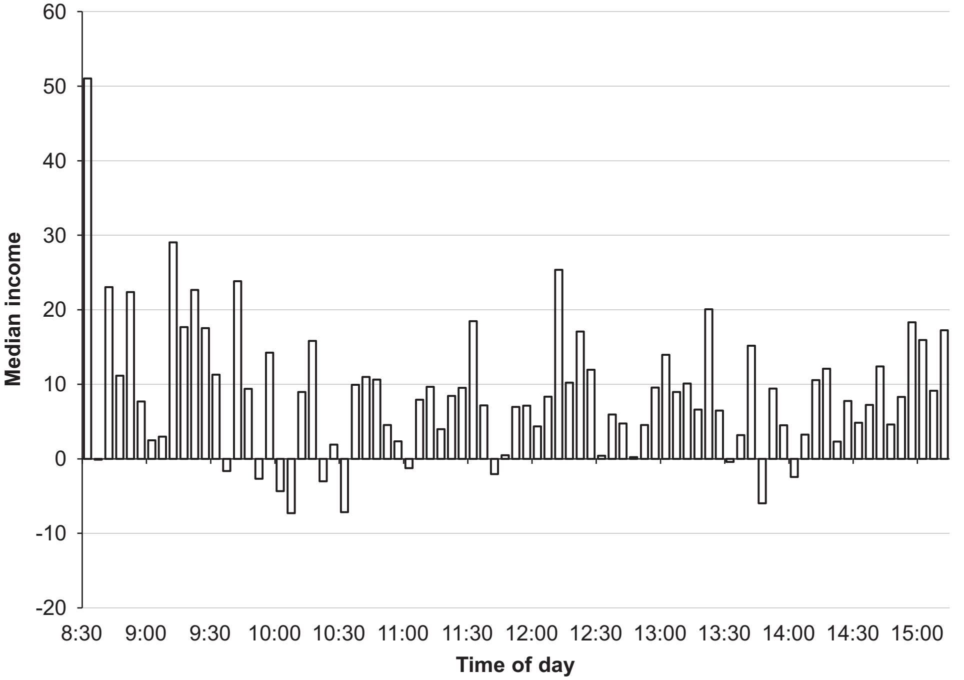

Fig. 3. Median intraday dealer income per contract. The data are S&P 500 futures trading from January 1998 through December 2001. Income is computed for all trades executed for floor traders' proprietary accounts within a five-minute. Income is calculated by marking each proprietary trade to market using the daily settlement price. Total income is then divided by the maximum of buy or sell trades within the interval to obtain a per contract measure. The median income in dollars is computed by each interval across the entire sample period.

undertaking. A lower frequency, such as hourly or daily, would lead to a greater proportion of offsetting trades, and thus more "spread" income.

Table 4 presents Pearson correlations for the returns based on customer buy prices, μ_t^a, order flow, and several income and volume measures. (Again, unreported results for μ_t^b, returns based on customer sell prices, are quite comparable.) All three measures of volume introduced in Table 1 are employed in this section: customer volume, hedger volume, and total volume. The correlation between the return, μ_t^a, and contemporaneous order flow is positive and statistically significant at 0.097, suggesting that, unconditionally, prices rise in response to a larger demand pressure from off-exchange customers.

We also find that order flow and volume are positively correlated, with net customer buying, on average, at times of high volume. There is also a large correlation between volume and income measures: unconditionally, dealers are earning larger income per contract when the level of trading activity is high. The correlation coefficient of 0.125 between customer volume and spread income suggests that, as expected, spread income is larger on average at times of low information asymmetry (high volume). In other words, when volume is high it is more likely (unconditionally) that a greater percentage of trades will be uninformed, so that *per trade* information is low, and spread income per contract is high.

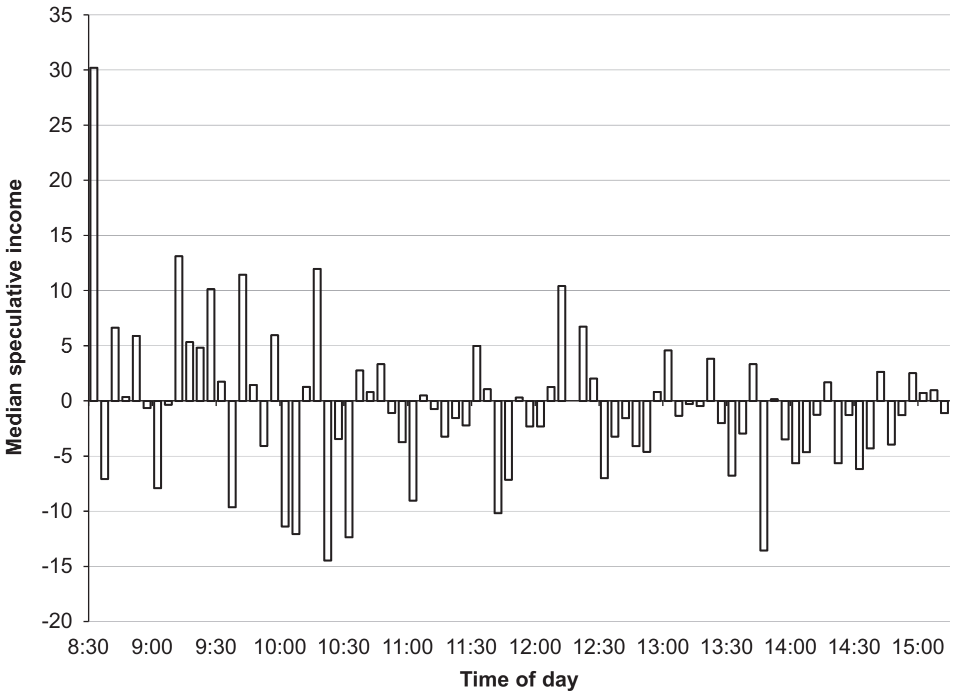

Fig. 4. Median intraday speculative income per contract. The data are S&P 500 futures trading from January 1998 through December 2001. Speculative income is calculated by marking the proprietary trade imbalance within a five-minute interval to market at the daily settlement. Total speculative income is divided by the absolute value of the trade imbalance to obtain a per contract measure. The median speculative income in dollars is computed by each interval across the entire sample period.

In order to capture the state dependency in the dynamics of the order flow polynomial, we use the logistic smooth transition (LSTR) framework, as in Teräsvirta (1994) and Van Dijk, Teräsvirta, and Franses (2000). However, before proceeding with the estimation of the LSTR model, we test for nonlinearity in the order flow and price relation and whether the LSTR framework is adequate for modeling the nonlinearity. Following Eitrheim and Teräsvirta (1996), we perform a diagnostic LM test that allows for nonlinearities potentially introduced by exogenous state variables. The results of these preliminary tests, available from the authors, strongly support nonlinear state dependence in the return dynamics and support adequacy of the LSTR framework for nonlinear estimation.

In our application, the LSTR model basically relates a vector of order flow variables, x_i, to the return μ_t^a, with this relation allowed to vary with the state variables. The fundamental specification, with Z_t representing any of the volume or income measures described above, takes the following form:

$$\mu_t^a = d_0(Z_t) + \sum_{i=-11}^{1} d_i(Z_t)x_{t-i} + w_t. \tag{9}$$

The functions $d_i(Z_t)$ describe state-varying parameters on the order flow variables x_i. The particular way that these parameters vary in response to changes in the state variable, Z_t, is

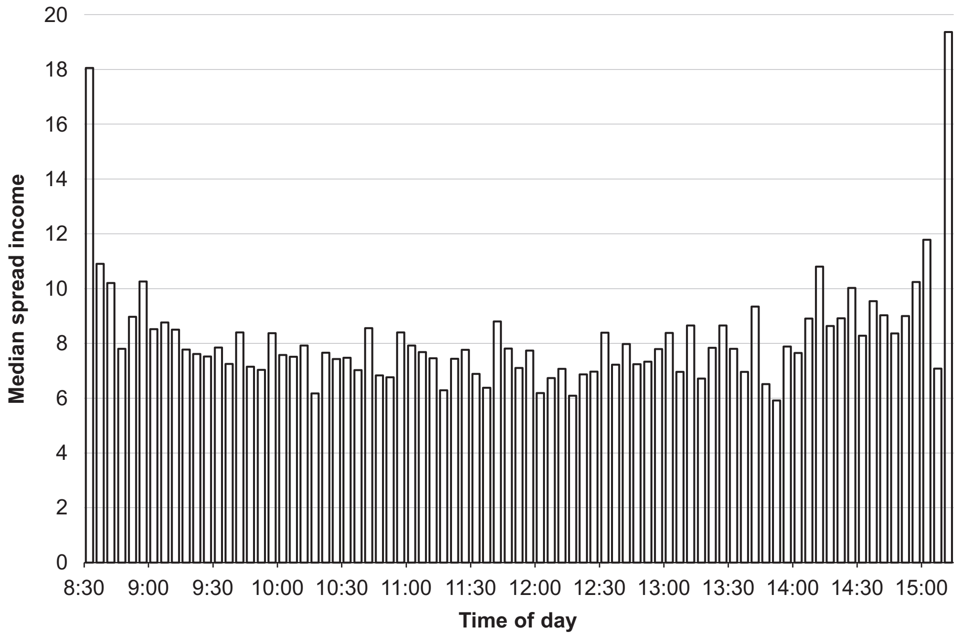

Fig. 5. Median intraday spread income per contract. The data are S&P 500 futures trading from January 1998 through December 2001. Spread income is the income (value of sales minus value of purchases) from offsetting trades (the minimum of buys and sells) for proprietary trading within a five-minute interval. Total spread income is divided by the minimum of buys and sells to obtain a per contract measure. The median spread income in dollars is computed by each interval across the entire sample period.

given by

$$d_i(Z_t) = d_i(-\infty)(1 - G(Z_t;\gamma,\zeta)) + d_i(\infty)G(Z_t;\gamma,\zeta),$$

with the function $G(Z_t;\gamma,\zeta)$ taking on the logistic smooth transition form

$$G(Z_t;\gamma,\zeta) = \frac{1}{1+\exp\{-\gamma(Z_t-\zeta)\}} \quad \text{where} \quad \gamma > 0.$$

The parameter γ plays the role of a transition parameter, providing the rate at which the coefficients d_i vary with the market interaction variable Z_t. The parameter value varies between the extremes of $d_i(-\infty)$, the parameter value given the smallest level of income or volume, and $d_i(-\infty)$, the parameter value given the highest level of income or volume. This can be compared to the bivariate model where the parameter values are fixed rather than variable. This specification allows for a smooth adjustment in the dependency of these parameters on the state variables. Starting values for γ and ζ are obtained as in Teräsvirta (1994) with details provided in the Appendix A.

Table 5 reports the estimates of the state-dependent LSTR model, where the order flow terms vary with market volume. For these volume measures we use lead volume, or Z_{t+1}, rather than contemporaneous volume, Z_t as the state variable. This choice is stipulated by the results of preliminary tests showing that Z_{t+1} captures a greater degree of nonlinearity in the relation between order flow and price. In view of the presence of serial dependence in

Table 4
Pearson correlation coefficients

				Income			Volume			
	x_t	x_{t+1}	x_{t-1}	Total	Spread	Speculative	Total	Customer	Hedger	μ^a_{t-1}
μ^a_t	0.097	−0.017	0.188	−0.013	0.021	−0.014	−0.013	0.006	−0.033	−0.096
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.083)	(0.000)	(0.000)
x_t		0.118	0.114	0.012	0.038	0.011	0.395	0.454	0.197	0.189
		(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)
x_{t+1}			0.130	0.012	0.043	0.011	0.251	0.279	0.137	0.098
			(0.000)	(0.000)	(0.000)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)
x_{t-1}				0.016	0.043	0.014	0.209	0.227	0.100	−0.024
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Income					0.061	0.999	0.019	0.022	0.010	−0.002
					(0.000)	(0.000)	(0.000)	(0.000)	(0.007)	(0.670)
Spread income						0.030	0.110	0.125	0.066	−0.001
						(0.000)	(0.000)	(0.000)	(0.000)	(0.880)
Spec. income								0.016	0.018	0.008
							(0.000)	(0.000)	(0.034)	(0.679)
Total volume								0.973	0.856	−0.005
								(0.000)	(0.000)	(0.218)
Cust. volume									0.727	0.000
									(0.000)	(0.950)
Hedger volume										−0.009
										(0.018)

Volume captures combined trading by all futures trader types. Customer volume is trading by off-floor customers. Hedger volume is trading for other exchange members on the floor, such as option traders using the futures to hedge. Income is calculated by marking each trade in an interval to market at the settlement price, and normalizing by the total number of contracts traded. Speculative income (per contract) is calculated by marking only the imbalance within an interval to market at the daily settlement. Spread income (per contract) is the income from offsetting trades within an interval. Order flow, x_t, is net customer buying in the interval. The return measure, μ^a_t, is the logarithmic change in customer purchase prices over a five-minute nterval *t-1* to *t*. Pearson correlations are computed for all dealer trades in S&P 500 futures contracts for the period from January 1998 through December 2001. *p*-values are reported in parentheses.

residuals of the model, we allow for first-order autocorrelation as well as heteroskedasticity. We compute the Newey-West version of nonlinear least squares standard errors.

The results suggest that in low-volume states, the transient impact of order flow is heightened. For example, for the customer volume specification, the coefficient on contemporaneous order flow, x_t, is 0.332 in the low-volume regime and −0.014 in the high volume regime. The result is similar when total volume is used as a state variable. Contemporaneous order flow has a more pronounced temporary effect on price when volume is low. Lead coefficients, that is, on x_{t+1}, are similar across volume regimes. If we consider the temporary effect of order flow on price to be captured jointly by coefficients on x_{t-1}, x_t, and x_{t+1}, this is clearly positive and greater when volume is low. Thus, in the short run, price is more sensitive to order flow when volume is low.

Quite the opposite occurs over the long run. The permanent effect, D^*, measured as the sum of the coefficients of the polynomial $D(L)$, is negative in the low-volume state (−0.15) and very large and positive in the high-volume state (0.33). The transition parameter γ,

Table 5
Estimates of nonlinear smooth transition regression with dealer volume interaction

	$D(L,Z_t)$ coefficients (x100,000)																
	Intercept	x_{t-1}	x_t	x_{t+1}	x_{t+2}	x_{t+3}	x_{t+4}	x_{t+5}	x_{t+6}	x_{t+7}	x_{t+8}	x_{t+9}	x_{t+10}	x_{t+11}	D^*	g*std(Z)	c
Baseline model																	
$D(L)$	−0.020	−0.025	0.139	0.264	−0.040	−0.070	−0.057	−0.050	−0.048	−0.046	−0.037	−0.028	−0.026	−0.015	−0.039		
	(0.491)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.000)		
Customer volume																	
$D(L,0)$	33.400	0.027	0.332	0.303	−0.063	−0.167	−0.103	−0.122	−0.068	−0.096	−0.060	−0.050	−0.053	−0.033	−0.153	0.95*498	1,105
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$D(L,+\infty)$	−0.480	−0.049	−0.014	0.297	0.025	0.035	0.016	0.018							0.328		
	(0.857)	(0.000)	(0.465)	(0.000)	(0.050)	(0.022)	(0.215)	(0.177)							(0.001)		
Total volume																	
$D(L,0)$	20.200	0.026	0.287	0.282	−0.063	−0.155	−0.098	−0.100	−0.064	−0.090	−0.053	−0.048	−0.048	−0.032	−0.156	1.26*728	1,388
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
$D(L,+\infty)$	3.700	−0.057	0.003	0.310	0.022	0.031	0.015	0.005							0.329		
	(0.055)	(0.000)	(0.873)	(0.000)	(0.075)	(0.033)	(0.242)	(0.696)							(0.001)		

The estimated model is $\mu_t^a = d_0(Z_{t+1}) + \sum_{i=-11}^{1} d_i(Z_{t+1})x_{t-i} + u_t$, with the polynomial lag coefficients given by $d_i(Z_t) = d_i(\infty)(1 - G(Z_t;\gamma,\zeta)) + d_i(-\infty)G(Z_t;\gamma,\zeta)$, and the transition function given by $G(Z_{t+1};\gamma,\zeta) = \frac{1}{1+\exp\{-\gamma(Z_{t+1}-\zeta)\}}$. The return measure μ_t^a is the logarithmic changes in customer purchase prices over a five-minute interval. The order flow, x_t, is net customer buying in the interval, and Z_t is the state variable, measures of volume, and γ and ζ parameters to be estimated. The starting values for NLS estimation are obtained using two-dimensional grid search results. The table presents the nonlinear least squares estimates and p-values are in parentheses. Standard errors are heteroskedasticity and autocorrelation consistent. The convergence algorithm used is Newton-Raphson. The permanent impact given by D^* is computed as the sum of the $D(L)$ polynomial coefficients.

when adjusted by the standard deviation of the market state variable, suggests a quick transition from low- to high-volume states.

The finding that the short-run sensitivity of price to order flow is higher when volume is low is intuitive. The effect of a sizeable imbalance in the chaotic, high-volume morning, for example, is probably hard to distinguish. In the low-volume midday, a similar imbalance could easily have a more obvious short-run effect. Of greater interest is the switch in the sign of the permanent effect from low- to high-volume regimes. One explanation for this is a switch in the volume mix, again easiest to discuss in terms of the time of day. For example, in the typically lower-volume midday, trades are more likely to be hedge related, coming from passive mutual funds, pension plans, and so forth. The negative permanent impact during this time can be interpreted as a cost or slippage associated with this hedging. In the higher-volume morning, the market is more likely to be digesting overnight information. Orders at this time are more likely to come from trades based on divergent opinions about public news. The net order flow from these trades has a positive long-run impact on price, so that the results of these speculations are, on average, correct.

Table 6 presents results of the LSTR model where two measures of floor trader income are used as the state variable that governs the regime switch. The income measures are speculative and total income, both analyzed in per-contract terms. (We do not report the results of the estimation of LSTR model with spread income since this variable does not exhibit the movement of the other income measures, which leads to a lack of convergence of the model.) Similar to volume, the contemporaneous price response coefficient is larger in the low-income state. For example, using total income, the x_t coefficient has a value of 0.262 in the low-income state, more than twice the value of 0.126 in the high-income state. Thus, in the short run, when income is low, price is more responsive to order flow. Note the interesting shift that occurs in the short-run effect. The coefficients associated with x_{t+1} are higher in the high-income regime (for example, for speculative income, the value is 0.275 in the high-income state vs. 0.097 in the low-income state). The aggregate short-run effect, the sum of the coefficients on x_{t-1}, x_t, and x_{t+1}, is 0.278 in the low-income state and 0.378 in the high-income state for speculative income. Similar results obtain for total income (0.286 vs. 0.379). Order flow has a greater transient impact when floor traders earn larger per-contract income.

The shift of the short-run impact from contemporaneous to lead order flow when floor trader income is higher is interesting. When income is low, the contemporaneous effect dominates, whereas when income is high, the lead effect dominates. The low-income effect is the traditional microstructure response, a simultaneous shift in bid and offers in response to an order imbalance. However, when income is higher, price adjusts in anticipation of the order flow, suggesting that sometimes the floor traders are able to shift price anticipating an order flow. This fits in well with the findings of Ferguson and Mann (2001) and Kurov (2005) that there is some information in the personal trading of floor traders. Based on our results, this semifundamental information is not constant, and could be in the form of fleeting inferences regarding pending order flow.

The results on the sum, D^*, the permanent effect of order flow on price, also reveal differences between the high-and low-income regimes. These differences are not nearly as dramatic as for the volume regimes. However, when income is low, the permanent order flow impact is near zero, with values of 0.005 for speculative income and 0.016 for total income. These compare to values of 0.023 and 0.024 for the high income regime. When traders are earning less (or losing), the order flow is hardly affecting price permanently.

Table 6
Estimates of nonlinear smooth transition regression with income interaction

	$D(L, Z_t)$ coefficients (x100,000)																
	Intercept	x_{t-1}	x_t	x_{t+1}	x_{t+2}	x_{t+3}	x_{t+4}	x_{t+5}	x_{t+6}	x_{t+7}	x_{t+8}	x_{t+9}	x_{t+10}	x_{t+11}	D^*	g∗std(Z)	c
Baseline model																	
$D(L)$	−0.020	−0.025	0.139	0.264	−0.040	−0.070	−0.057	−0.050	−0.048	−0.046	−0.037	−0.028	−0.026	−0.015	−0.039		
	(0.491)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)	(0.000)		
Speculative income																	
$D(L,-\infty)$	−26.000	−0.081	0.262	0.097	0.089	−0.125	−0.087			−0.091	−0.059				0.005	4.07∗1,498	3,662
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			(0.000)	(0.019)				(0.000)	(0.000)	(0.000)
D$(L,+\infty)$	−2.000	−0.023	0.126	0.275	−0.042	−0.066	−0.047	−0.051	−0.033	−0.035	−0.023	−0.021	−0.021	−0.016	0.023		
	(0.025)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)		
Total income																	
$D(L,-\infty)$	−25.000	−0.069	0.241	0.114	0.065	−0.122	−0.073			−0.086	−0.054				0.016	6.84∗1499	3,267
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.009)			(0.000)	(0.021)				(0.000)	(0.000)	(0.000)
$D(L,+\infty)$	−2.000	−0.023	0.127	0.275	−0.042	−0.065	−0.047	−0.051	−0.033	−0.034	−0.024	−0.022	−0.021	−0.016	0.024		
	(0.025)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)		

The estimated model is $\mu_t^a = d_0(Z_{t+1}) + \sum_{i=-11}^{1} d_i(Z_{t+1})x_{t-i} + u_t$, with the polynomial lag coefficients given by $d_i(Z_t) = d_i(\infty)(1 - G(Z_t;\gamma,\zeta)) + d_i(-\infty)G(Z_t;\gamma,\zeta)$, and transition function given by $G(Z_{t+1};\gamma,\zeta) = \frac{1}{1+\exp\{-\gamma(Z_{t+1}-\zeta)\}}$. The return measure μ_t^a is the logarithmic changes in floor trader sales prices over a five-minute interval. The order flow, x_t, is net customer buying in the interval, Z_t is the state variable, one of the two income measures, and γ and ζ are parameters to be estimated. Floor trader income, total and speculative, is normalized to obtain income per contract. Starting values are obtained using two-dimensional grid search results. The table presents the non-linear least squares estimates and p-values are in parentheses. Standard errors are heteroskedasticity and autocorrelation consistent. The convergence algorithm used is Newton-Raphson. The permanent impact given by D^* is computed as the sum of the $D(L)$ polynomial coefficients.

Instead, floor traders profit when the permanent effect is slightly positive. One interpretation is that the low-income permanent effect is associated with times of predominant hedging, and the high-income effect is associated with a more informed order flow. This would suggest that floor traders do not profit (much) from simply processing hedger trades. Instead, the rarer times when the price discovery process is permanently and positively affected by net order flow corresponds to the time when floor traders earn greater income.

5. Conclusions

We analyze price discovery on the floor of the futures exchange, looking at the role of the futures floor trader as a market maker who makes inferences regarding price from order flow. We find that there is an equilibrium distribution of transaction prices across these traders on the futures floor at any point of time. The variability in price across these traders is small relative to fundamental price volatility, a finding substantially different than FX studies such as Evans (2002). Order flow, the net trading of off-exchange customers, has differing short-run and long-run price effects. The short-run effect is generally positive, as expected, with net customer buying having an immediate positive impact on price. The long-run effect is much smaller and, while significant, on average it is negative, suggesting a slight hedging motive, with some costly slippage as a result, for futures customers' trades. The finding of these two effects complements the work of Ferguson and Mann (2001) who speculate that futures floor traders could infer semifundamental private information due to their physical location. Our results serve to confirm this through the finding that, by far, the largest impact of the futures order flow on price is only temporary. Thus, similar to the FX studies, such as Ito, Lyons, and Melvin (1998), there is at most only semifundamental information obtainable from the futures order flow.

In addition, we find that both the long-and short-run price sensitivity of the futures order flow vary depending on the state of the market, which we measure by volume and floor trader income. This corresponds to the finding in Evans (2002) that the informational content of the FX order flow varies with the market state. Using a smooth transition methodology, we find that the futures order flow has a greater short-run impact both in low-volume and in low floor trader income states. There are also changes in the long-run price impact of order flow across volume and income regimes. When volume is high, the long-run impact is positive, so that informed trading seems to be occurring in high-volume periods, such as at the morning opening. The impact is also positive when floor trader income is high. The floor traders are not earning much income when the permanent impact is low. We conclude that the floor traders do not profit much from processing hedger orders, and earn more income during times of greater price discovery, when the long-run impact of order flow is significantly positive.

Finally, we find little long-run price impact of the S&P 500 futures order flow, on average, and yet previous studies have typically found a significant lead in the S&P 500 futures with respect to the S&P 500 cash index. An interesting extension would be to relate the futures *basis* to order flow, both futures and cash, or more broadly to incorporate cross market order flow measures into a model of several cointegrated prices. Further, order flow related to *arbitrage*, or quasi-arbitrage, associated with unusual values for the basis, could be distinguished from the impact of order flow during normal basis times.

Appendix A. Estimation Procedures

A.1. State space estimation

For estimation we write the model in state-space form and obtain GMM estimates of the coefficients. The state-space form representation is given by

$$\psi_t = A\psi_{t-1} + \xi_t, \tag{A.1}$$

$$y_t = H\psi_t,$$

where ψ_t is a q-dimensioned state vector, and y_t is an r-dimensioned vector of observed variables, i.e., lead and lagged order flow terms from Eq (5). The matrix A represents the autocorrelation structure of the states, ψ_t. The matrix H transforms the states into observed variables. The term ξ_t is a q-dimensioned vector of shocks with zero means that are serially uncorrelated and have covariance matrix Ω. Using (A.1), the covariance of the states, $\Gamma(k) = Cov(\psi_t, \psi'_{t-k})$, is computed as $\Gamma(k) = A\Gamma(k-1)$ with $\Gamma(k) = vec^{-1}[(I - A \otimes A)^{-1} vec(\Omega)]$.

The covariance of the observed variables is therefore given by

$$Cov(y_t, y'_{t-k}) \equiv \gamma(k) = H\Gamma(k)H'. \tag{A.2}$$

The orthogonality conditions are as in the standard GMM case and have the following form:

$$\mathrm{E}[m_t(k;\theta)] = 0, \tag{A.3}$$

$$\text{where} \quad m_t(k;\theta) = J(k)\begin{bmatrix} vec(y_t z'_{t-k}) \\ vec(y_t y'_{t-k} - \gamma(k;\theta)) \end{bmatrix}$$

for $k = 1,2,\ldots,K$. $J(k)$ is a vector of ones and zeros that selects the moments to be included in the form $m_t(k;\theta)$. To compute the GMM estimates we select a vector of moment conditions and find the parameter values θ that minimize $Q(\theta) = m_T(\theta)' W^{-1} m_T(\theta)$, where $m_T(\theta) = \frac{1}{T}\sum_1^T m_t(\theta)m$. The weighting matrix W follows Newey and West (1987) of the first order. Efficient estimates are computed by using a two-step GMM procedure.

A.2. LSTR starting values

We perform a two-dimensional grid search to obtain starting values for γ and ζ. As suggested by Teräsvirta (1994), the interaction variable Z_t is normalized by dividing it by its standard deviation. Initially, we estimate the OLS coefficient d_i as a function of parameters γ and ζ and then run a two-dimensional grid search over 100 possible values of γ (from 0.1 to 10) and 100 possible values of ζ (from 100 to 10,000) to find the values of parameters that minimize the function $Q_T(\gamma,\zeta)$, with respect to two parameters γ and ζ:

$$Q_T(\gamma,\zeta) = \sum_{t=1}^{T} (\mu_t^a - \delta(\gamma,\zeta)'\xi_t(\gamma,\zeta))^2, \tag{A.4}$$

where

$$\xi_t(\gamma,\zeta) = \{(1-G_t), G_t, x_{t-1}(1-G_t), x_{t-1}G_t, x_t(1-G_t), x_tG_t, x_{t+1}(1-G_t), x_{t+1}G_t\}$$

with

$$G(Z_t;\gamma,\zeta)=\frac{1}{1+\exp\{-\gamma(Z_t-\zeta)\}},$$

and

$$\delta(\gamma,\zeta)=\left(\sum_{t=1}^{T}\xi_t(\gamma,\zeta)\xi_t(\gamma,\zeta)'\right)^{-1}\left(\sum_{t=1}^{T}\xi_t(\gamma,\zeta)\mu_t^a\right).$$

The grid search offers reasonable starting values for the nonlinear estimation. Next, we compute OLS coefficients based on the LSTR parameters γ and ζ and use these estimates as starting values for nonlinear least-squares estimation.

References

Admati, A., Pfleiderer, P., 1988. A theory of intraday patterns: volume and price variability. Review of Financial Studies 1, 3–40.

Cao, H., Evans, M., Lyons, R., 2006. Inventory information. The Journal of Business 79, 325–364.

Chae, J., 2005. Trading volume, information asymmetry, and timing information. Journal of Finance 60, 413–442.

Chakravarty, S., Li, K., 2003. An examination of own account trading by dual traders in futures markets. Journal of Financial Economics 69, 375–397.

Chordia, T., Roll, R., Subrahmanyam, A., 2002. Order imbalance, liquidity and market returns. Journal of Financial Economics 65, 111–130.

Cumby, R., Huizinga, J., 1992. Testing the autocorrelation structure of disturbances in ordinary least squares and instrumental variables regressions. Econometrica 60, 185–195.

Easley, D., O'Hara, M., 1992. Time and the process of security price adjustment. Journal of Finance 47, 577–605.

Eitrheim, O., Teräsvirta, T., 1996. Testing the adequacy of smooth transition autoregressive models. Journal of Econometrics 74, 59–75.

Evans, M., 2002. FX trading and exchange rate dynamics. Journal of Finance 57, 2405–2447.

Evans, M., Lyons, R., 2002. How is macro news transmitted to exchange rates? Working paper. Georgetown University, Washington, DC.

Ferguson, M., Mann, S., 2001. Intraday variation in futures markets. Journal of Business 74, 125–160.

Ferguson, M., Mann S., Weisburd, A., 2004. Liquidity, anonymity, and competitive market makers. Working paper. Neeley School of Business, Texas Christian University, Fort Worth, TX.

Foster, D., Viswanathan, S., 1990. A theory of intraday variations in volumes, variances, and trading costs. Review of Financial Studies 3, 593–624.

Glosten, L., Milgrom, P., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 17, 99–117.

Granger, C., Teräsvirta, T., 1993. Modelling Nonlinear Economic Relationships. Oxford University Press, Oxford, UK.

Ito, J., Lyons, R., Melvin, M., 1998. Is there private information in the FX market? The Tokyo Experiment. Journal of Finance 80, 1111–1130.

Kraus, A., Stoll, H., 1972. Parallel trading by institutional investors. Journal of Financial and Quantitative Analysis 7, 2107–2138.

Kurov, A., 2005. Execution quality in open-outcry futures markets. Journal of Futures Markets 25, 1067–1092.

Kurov, A., Lasser, D., 2004. Price dynamics in the regular and E-mini futures markets. Journal of Financial and Quantitative Analysis 39, 365–384.

Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Lakonishok, J., Shleifer, A., Vishny, R., 1992. The impact of institutional trading on stock prices. Journal of Financial Economics 32, 23–43.

Lee, C., 1992. Earnings news and small traders: an intraday analysis. Journal of Accounting and Economics 15, 265–302.

Luo, J., 2001. Market conditions, order flow and exchange rates determination. Working paper. University of London-Financial Markets Group, London.
Manaster, S., Mann, S., 1996. Life in the pits: competitive market making and inventory control. Review of Financial Studies 9, 953–975.
Manaster, S., Mann, S., 1999. Sources of market making profits: man does not live by spread alone. Working paper. Neeley School of Business, Texas Christian University, Fort Worth, TX.
Newey, W., West, K., 1987. A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Ready, M., 1999. The specialist's discretion: stopped orders and price improvement. Review of Financial Studies 12, 1075–1112.
Stoll, H., 2000. Friction. Journal of Finance 55, 1479–1514.
Teräsvirta, T., 1994. Specification, estimation, and evaluation of smooth transition autoregressive models. Journal of the American Statistical Association 89, 208–218.
Van Dijk, D., Teräsvirta, T., Franses, P., 2000. Smooth transition autoregressive models—a survey of recent developments. Econometric Institute Research Report EI2000-23/A.
Wermers, R., 1999. Mutual fund herding and the impact on stock prices. Journal of Finance 54, 581–622.
Working, H., 1953. Hedging reconsidered. Journal of Farm Economics 35, 544–561.