Divergent Expectations¹

When investors agree to disagree

Paul L. Davis
Retired Senior Managing Director
TIAA-CREF Investment Management LLC

Michael S. Pagano
Associate Professor of Finance
College of Commerce and Finance,
Villanova University

Robert A. Schwartz
Marvin M. Speiser Professor of Finance
Zicklin School of Business
Baruch College, CUNY

Journal of Portfolio Management, Volume 34, Number 1, Fall 2007, pp. 84-95.

Draft: January 27, 2007
Please do not quote or distribute without permission of the authors

¹ Parts of this paper draw from some sections of “Markets in Transition: Looking Back and Peering Forward,” by the same authors, which is the first chapter of a book titled Banken, Börsen und Kapitalmärkte (Banks, Exchanges, and Capital Markets), Wolfgang Bessler, editor, 2006. We thank Wolfgang Bessler for his comments on an earlier draft of this paper.
Divergent Expectations

When investors agree to disagree

Abstract

Investors who possess the same information and interpret it differently are said to have divergent (as distinct from) homogeneous expectations. Financial economists have widely frowned on the divergent expectations assumption. Nevertheless, this assumption describes reality and is critically important. It paves the way to understanding price and quantity discovery as major functions of a marketplace, and it goes to the heart of an important question – what drives trading and why does market structure matter? Many issues concerning market structure and market structure regulation should be analyzed in a divergent expectations context. This paper considers what is involved in terms of market participants and public policy.
Divergent Expectations

In a piece first published in Latin in 1738, the mathematician and scientist, Daniel Bernoulli wrote, "...the determination of the value of an item must not be based on its price, but rather on the utility it yields. The price of the item is dependent only on the thing itself and is equal for everyone; the utility, however, is dependent on the particular circumstances of the person making the estimate."\(^2\) The statement is germane to the topic of this paper if one substitutes "expected return and risk" for "utility," and "information set" for "the particular circumstances..." This is because an equity’s share price is the same for everyone, but each person forms his or her own expectations about the stock’s future expected return and risk based on a common information set that they may all possess.

Investors who possess the same information and interpret it differently are said to have divergent expectations. Financial economists involved with security pricing have generally frowned on using this assumption, and have alternatively taken the expectations of similarly informed investors to be homogeneous. As implausible as it may be, it is understandable why the profession has taken homogeneous expectations as a cornerstone assumption for much theoretical and empirical analysis – the simplification greatly facilitates the mathematical modeling of financial problems. Furthermore, some academicians consider homogeneity reasonable. After all, give rational decision makers the same information and they should, the thinking goes, reach identical conclusions.

The homogeneous expectations assumption may be a reasonable modeling device for applications concerning long run inter-stock pricing relationships such as those provided by, e.g., the Capital Asset Pricing Model and Arbitrage Pricing Theory. But reality does not conform to the homogeneous expectations assumption, and many issues concerning market structure and market structure regulation should be analyzed in a divergent (or heterogeneous) expectations context. Our objective in this paper is to consider what this means and what is involved for market participants.

Three decades ago, Miller (1977) took issue with the homogeneous expectations assumption. He wrote, "However, it is implausible to assume that although the future is...

\(^2\) See Bernoulli (1954).
very uncertain, and the forecasts are very difficult to make, that somehow everyone makes identical estimates of the return and risk from every security. In practice, the very concept of uncertainty implies that reasonable men may differ in their forecasts.” In his analysis of the divergent expectations environment, Miller focuses on longer run pricing implications. Interestingly, he shows that stocks in the highest risk classes can, over time, yield very low returns. In contrast, we focus our attention on shorter run insights concerning market structure, trading, and price discovery.

Despite having experienced over three decades of major structural change in equity markets around the globe, many questions concerning market structure and market structure regulation remain unanswered. Regulatory policy has paid major attention to items such as commission rates, bid-ask spreads, transparency and tick sizes that are important regardless of the expectations environment. But regulators have paid scant attention to the accuracy of price discovery and the completeness of quantity (size) discovery, two issues that are of major importance in a divergent expectations environment but which in a homogeneous expectations environment do not have particular prominence.

On the academic front, more attention has recently been given to analyzing a divergent expectations environment. Nevertheless, the assumption that investor expectations are homogeneous continues to underlie much market microstructure research on issues ranging from the components of the bid-ask spread and the informational content of price changes, to the strategic decisions informed traders make to exploit their edge when they possess information that is not available to all. An important example of the homogeneous expectations paradigm is Milgrom and Stokey’s (1982) well-known “no-trade” theorem. They showed that, in the absence of liquidity traders, no transaction will ever be made even if some investors possess new information that would lead them to want to trade with others who do not have this information.

---

3 A number of papers have established that the heterogeneity of expectations can affect market equilibrium prices. See, for instance, Miller (1977), Foucault (1999), Handa, Schwartz and Tiwari (2003), Paroush, Schwartz and Wolf (2007), and Scheinkman and Xiong (2003).

4 Liquidity traders are investors who are buying or selling securities to satisfy idiosyncratic cash flow needs (e.g., investing a bonus in a stock or selling securities to make a down payment on the purchase of a house).
The intuition behind the Milgrom-Stokey result is that any uninformed participant, upon observing the arrival of someone looking, e.g., to buy, knows that, with no liquidity traders present, the new arrival must possess information that would justify shares being valued at a higher price. Thus, the uninformed seller will not accept any price that the informed trader would be willing to pay. In a Milgrom-Stokey world, all participants (both informed and uninformed) know what the others are thinking, and everyone knows that everyone else knows what everyone is thinking. Consequently, the uninformed will not trade with the informed and the market will not be viable.

In this rational expectations setting, Grossman and Stiglitz (1976) have shown that, to be viable, an equity market must comprise three types of participants: information traders, liquidity traders, and “noise” traders. Black (1986) further examined the importance of noise trading, which he identified as trading on the basis of short-term random fluctuations, or noise, as if it is meaningful information. Momentum players and other market technicians are good examples of Black-type noise traders. Implicit in the classical trichotomy of information, liquidity, and noise trading is the assumption that all traders in possession of the same information agree on the import of that information. Thus, the expectations of the informed are homogeneous, they translate information into identical share values, and shares have unique values that are determined by the information.

How well does this structure describe real world equity markets? How valuable is the guidance that it gives to market structure development and to market structure regulation? Equity markets are indeed complex institutions; do we know enough about their operations to be able to design them and to regulate their structure without running afoul of unintended consequences? We suggest that reliance on a homogeneous expectations-based approach has led to regulation that is best-suited for retail investors (whose trading is commonly motivated by liquidity needs) but has hindered the ability of information-motivated institutional investors to trade in large quantities, thus decreasing the market’s overall efficiency. Another relevant question is: how far have we come in recent years? We address these and related questions in this paper. We start by taking a brief look at where we are today.
How Far Have We Come?

In recent decades, technological change and regulatory initiatives have thoroughly reshaped the equity markets, the organizational structure of exchanges, the competitive environment, and the operations of equity traders. Trading volumes have soared, commissions have shrunk and, although they are difficult to measure, the implicit costs of trading have been lowered, at least in the opinion of many market professionals. As the twentieth century turned into the twenty first, a new breed of traders has emerged. The young bucks at the trading desks of today have grown up with computer technology. Unlike their predecessors, they are comfortable with the speed and control that a computer can deliver, and they welcome the electronic technology that today is pervasive on buy-side trading desks, in broker/dealer firms, on exchanges, and throughout the alternative trading systems (ATSs). Surveying these developments, a student of the market might conclude that great progress has been made. It might appear that we have gone much of the way to the Promised Land.

Not quite. Over the past thirty years, many questions have been raised and issues debated, but the debates continue unabated and important questions remain unanswered. True, today, small orders for large cap blue chips present no problem; they can be executed with lightning speed at current prices. But big orders for stocks of all market cap sizes and all orders for mid and small cap stocks are not easily dealt with. New block trading facilities like Liquidnet and Pipeline have come into being and are attracting appreciable order flow, but they currently account for only a small fraction of total trading volume. Electronic call auctions are being widely used to open and close markets, but their potential for handling institutional order flow has not been adequately exploited. Regulators on both sides of the Atlantic call for transparency but, while they might require that certain orders be exposed, it ultimately is not possible to legislate transparency. Markets have consolidated on both sides of the Atlantic and the trend will no doubt continue, but regulatory authorities fear the advent of monopoly power even as new players such as Liquidnet and Pipeline emerge.

There are further issues. Best execution, a regulatory requirement in the U.S. that was first set forth in the 1975 Securities Acts Amendments, is now slated for the European arena with the coming implementation of the Markets in Financial Instruments
Directive (MiFID), but best execution is virtually impossible to define and very difficult to implement. Soft dollar arrangements remain unsavory in the opinion of many, but a forced unbundling of trading, research, and other soft dollar products can have its own undesirable consequences. With computerized order handling and trade execution, speed is now measured in milli-seconds, but trading remains a zero sum game and it is difficult to understand how traders, as a community, can benefit from operating in a continuous market environment with sub-second response times. Would we all be better off if everybody used Formula-one racing cars for commuting and daily shopping trips?

While markets to some extent may have consolidated spatially, they are fracturing temporally as large orders are increasingly being sliced and diced and fed to the markets one small tranche at a time. NYSE average trade size in 1988 was 2,303 shares; by June, 2005, it had fallen to 343 shares. NYSE block-trading in 1988 was 51.1 percent of reported volume; in June, 2005, it had declined to 27.9 percent. Trading information in general, and quotes in particular, can change with startling rapidity in fast, computerized markets, which means that traders need to have new tools at their disposal. Accordingly, in the past couple of years, a great deal of attention has turned to algorithmic trading (also referred to as “algos”). However, one might question whether the widespread use of algos represents a valuable step forward for traders as a group, or is a symptom of continuing friction, illiquidity, high market impact costs, accentuated intra-day volatility, and fuzzy pricing in our markets.

What Motivates Trading?

A critical starting point for thinking about market microstructure resides in the complexities of a deceptively simple question: what motivates individuals to trade? Alternatively stated, how good can our understanding of the markets be if we cannot properly answer this very basic question?5

As we have noted, the classical academic answer is that there are three types of traders – information, liquidity, and noise traders, with informed traders having the same (homogeneous) expectations about the returns that a stock will deliver. Four major conclusions follow from the homogeneity assumption:

5 For further discussion of the motives for trading, see Sarkar and Schwartz (2007).
• Information maps uniquely into security values. Alternatively stated, securities have intrinsic (i.e., true or fundamental) values that are uniquely linked to the information set.

• Stocks prices that deviate from intrinsic values will be driven back to their fundamentals by arbitrageurs. Arbitrage is a key mechanism in both the Capital Asset Pricing Model (CAPM) and Arbitrage Pricing Theory (APT).

• If all participants are “informed” (i.e., know the full information set) and trades are triggered for liquidity reasons only, shares will trade at bid and ask quotes that are appropriate given the intrinsic values and, aside from the bid-ask bounce, prices will follow random walks with drift.

• Informed participants will never trade with each other because they have the same expectations; consequently, liquidity and noise traders must be present for a market to be viable.

But homogeneity is only an assumption, and the four bulleted points fit poorly with reality. Accordingly, academic belief that investor expectations based on a common information set are homogeneous has been tempered of late. It is recognized that some participants produce “private information” (namely, that they further process information so as to gain insights that are not immediately apparent).6 One striking illustration was reported in the Wall Street Journal on October 6, 2003. As summarized by Schwartz, Francioni, and Weber (2006),

“The Journal reported that between June and August 2003 certain large institutional investors, using two detailed pharmaceutical databases that cost subscribers between $25,000 and $50,000 a year, were able to determine that the drug company Schering-Plough was losing market share for its hepatitis C medicine to a competitor, Roche Holding AG. The investing public got hints of the development in July, but full information was not available until August 22. During this period, Schering’s share price peaked at $20.47 on June 18 and then drifted down, dipping below $17 as July turned into August… on August 22, the day when Schering confirmed the declining sales trend for its product, shares opened down $2.18 at $14.30.”

6 Also see footnote 3.
Whether participant expectations differ because of the actual “production” of private information as in the Schering-Plough case, or simply because different people interpret information differently, the expectations of a group of investors can be divergent. A spectrum of market realities calls attention to the importance of understanding this reality:

- Information sets are vast and enormously complex, and our tools for analyzing them are relatively crude. Consequently, precise assessments and translations into exact share values are not possible. Imprecision on this front leads to disagreements about share value.

- Analysts’ recommendations commonly differ.

- Much short selling is no doubt triggered by participants disagreeing with market assessments of share values.

- Institutional investors commonly execute large trades (e.g., 100,000 shares or more) in individual stocks with each other. Trading in new ATSs such as Liquidnet or Pipeline offers compelling evidence of large institution-to-institution trades. These trades are not likely to be motivated by liquidity needs (lists or basket trades would be used for that purpose). The participants are also unlikely to be noise traders. Thus, they must be information traders who disagree with one another about share value.

With divergent expectations, trades are not made simply because some participants have superior information while others are experiencing idiosyncratic cash flows (the liquidity traders), and yet others believe that they have detected exploitable patterns in the data (the noise traders). With divergent expectations, participants may trade with each other simply because they disagree with each other. Such trading may occur as a natural part of the dynamic price discovery that occurs following any news release, and it can explain the volume surges that have been observed to follow news releases.7

If participants are free to disagree with each other, then they are free to change their expectations for their own individual reasons. This reality sheds a different meaning on the term “liquidity trading,” as well as on the dichotomy between liquidity and

---

7 For further discussion, see Paroush, Schwartz, and Wolf (2007).
information trading. Furthermore, along with being free to change their expectations idiosyncratically, participants may also be influenced by what they observe others are thinking and doing. Paroush, Schwartz and Wolf (hereafter PSW, 2007) refer to this as “adaptive valuations,” the topic to which we next turn.

From Divergent Expectations to Adaptive Valuations

Once divergent expectations are allowed for, a further reality of the marketplace can be recognized. Namely, if individuals can form different expectations when faced with common information, then each may be influenced by the decisions of others. That is, they may have “adaptive valuations.” PSW have modeled trading and price discovery in an environment that is characterized by both divergent expectations and adaptive valuations. In so doing, they assume that any individual’s assessment of share value is a function of two things: (i) the publicly available information set, and (ii) the aggregate assessment of others (represented in PSW by the proportion of participants who are relatively bullish).

With divergent expectations and adaptive valuations, we have a situation where participants will trade with each other in the absence of information advantages and liquidity considerations. Further, the very process of trading affects participant assessments and share values. In this environment, as PSW show, shares do not have unique equilibrium values (and they certainly do not have unique intrinsic values). Rather, we are now in a multiple-equilibria framework within which price discovery is a path-dependent process. In such an environment, the behavioral responses of participants acquire importance.

---

8 Note that asymmetric information can still be present within a divergent expectations framework. That is, some traders might actually have superior information relative to other traders, and they will be able to trade with less-informed liquidity and noise traders, as well as with each other, if their expectations based on the superior information are divergent. In this sense, the traditional trichotomy of informed, liquidity, and noise traders can be viewed as a special case of the divergent expectations framework.

9 The literature on “information cascades” has also recognized that investors can have adaptive valuations. For example, if several buyers consecutively enter a market and bid up a security’s price, then others observing this activity might follow suit and thus form an information cascade (i.e., the trading actions of earlier traders “cascades” to later traders). This behavior can lead to speculative bubbles and crashes when there are divergent expectations about other traders’ valuations. For further details on information cascades, see Anderson and Holt (1996, 1997) and Ball and Holt (1998).
The process can be visualized with the use of an illustration. Assume that 100 participants are each placing a bet on the number of jelly beans in a jar filled with 2,500 beans. Each participant views the jar individually, and then each steps forward, one at a time in random order, to state a guess that is then revealed to everybody. As the betting precedes, participants are free to adjust their estimates based upon what they learn the others are guessing. Consider the bet that would be place by, e.g., a 31st participant when his or her initial guess is 1,800 beans and the average bet of the preceding 30 participants is 2,700 beans. We expect that the 31st participant’s bet will be a weighted average of 1,800 (the original guess) and 2,700 (the average guess of 30 others).

In this illustration, the large jar of 2,500 beans represents an information set that, because of its size, is difficult for each individual to assess with precision; the bets of the first 30 participants represents price discovery inputs to that point; and the response of the 31st participant to the collective valuations of the first thirty represents adaptive valuations. It can be shown that, as the betting progresses, the collective guess will eventually converge on a value that we might refer to as an equilibrium or consensus value. Interestingly, if the first participants to arrive happen by chance to place low bets, the equilibrium converged on will be relatively low. Conversely, if they happen by chance to place high bets, the equilibrium converged on will be relatively high. Accordingly, the game is a multi-equilibria, path-dependent process.

The multiple equilibria, path-dependent environment just set forth differs appreciably from a classical marketplace comprised of informed, liquidity, and noise traders; characterized by homogeneous expectations among the informed; and described as a unique equilibrium process that is not path-dependent. In the divergent expectations, adaptive valuation environment:

- Information sets do not map into unique security valuations in a deterministic manner that is independent of market processes.
- Price discovery (e.g., finding an equilibrium or consensus value) is a complex, dynamic process; this is not the case in a homogeneous expectations environment where informed participants know true values.
- Elevated short term (i.e., intraday) price volatility can be more richly understood.
• Tools related to technical analysis and algorithmic trading can acquire value.

• Achieving acceptably accurate price discovery becomes an important public policy issue, along with lowering commissions and shrinking the size of bid-ask spreads.

• The clarity of information release has heightened value, along with the speed and fairness of information disclosure.

• The seeds of longer-period bubbles and crashes may be found in the dynamic, intra-day process of price discovery.\(^{10}\)

• Price discovery has a psychological / behavioral component (i.e., what signals do participants look at to infer what each other is thinking, and how do they respond to these signals?).

• Noisy price discovery is associated with incomplete quantity discovery (i.e., the difficulty that large trading counterparties have in finding each other and trading the total number of shares that they are looking either to buy or to sell).

In sum, replacing the fundamental, simplifying assumption of homogeneous expectations with the more realistic assumption that participant expectations are divergent and adaptive can profoundly impact our thinking about how markets operate. Further investigation of the divergent expectations, adaptive valuations environment should shed useful new light on a spectrum of market structure and public policy issues that continue to remain unresolved.

**A Buy-side Perspective on Quantity Discovery**

Assets under management run the gamut from a few thousand dollars that are under the control of one individual, to a few million dollars that are handled by a small fund, to multi-billions of dollars that are managed by a large institution. As a consequence, and recognizing the existence of small retail orders and large orders that are commonly sliced and diced, an equity market’s network must be flexible enough to

---

\(^{10}\) See Paroush, Schwartz and Wolf (2007) for further discussion.
accommodate orders ranging from 100 shares to 1,000,000 shares or more. With regard to 100 share orders, a market maker can set bids and offers, trade across the spread, and make money by capturing the spread, as has been analyzed extensively in the academic literature. A market maker will trade 100 shares all of the time, 1,000 shares most of the time, and 10,000 shares some of the time. We are in a different world when really large orders (100,000 shares and more) are entered in the marketplace.

There are two fundamentally different approaches to trading a large order (for example, a 1,000,000 share buy order). First, a trader may slice and dice the order and feed it to the market one small tranche at a time. This approach is harmonious with the fundamental design of an order driven electronic market – a so-called “fast market.” However, as is pointed out by Davis, Pagano, and Schwartz (2006), when an order is sliced and diced, its entire execution may take quite a while. In other words, while the individual tranches are executed speedily in a fast market, the full order itself is executed slowly. Does an institutional investor have an alternative to slicing and dicing?

The second approach is to use the full search capabilities of the market network to find a natural seller to take the other side of the transaction. If participant expectations are divergent, the search is more likely to be successful. The seller may also be slicing and dicing. Or, the seller may not yet be active in the market, but can be planning to enter soon if conditions are suitable. Knowing that there is a buyer available can give the seller an incentive to step into the market and trade. Or, perhaps the seller is not in the market because he or she does not want to sell at the present market price, but would be willing to sell at a premium. The buyer may be willing to pay the premium so as to execute the entire order speedily. A seller might also be found at an ATS like Liquidnet or Pipeline, at a broker’s trading desk (using either dealer capital or the broker’s knowledge of potential contras), at a crossing network, or at an exchange’s call auction. In recent articles, searching for, and finding, the contra-side of a large order has been called “quantity discovery.”

---

11 The full “market network” is much more than the stock exchange. It includes brokers who, along with being the primary exchange’s customers, have trading desks that compete with stock exchanges. Alternative Trading Systems are also part of the market network.

12 For a further discussion of quantity discovery, see Davis, Pagano, and Schwartz (2006); Chakraborty, Pagano, and Schwartz (2006); and Sarkar and Schwartz (2007).
Ideally, quantity and price discovery should go hand-in-hand, as they do in traditional classroom presentations of demand and supply analysis. The structure of a marketplace is an often overlooked factor with regard to achieving an ideal price and quantity solution. An electronic order driven market – the central marketplace in most countries – must be augmented by other market mechanisms so as to meet the needs of large buyers and sellers. As additional mechanisms come into play, price discovery and quantity discovery usually de-couple. As they do, price discovery typically occurs at the central marketplace, quantity discovery gravitates to a facility more suitable for institutional trading, and the off-board (or off-book) facility typically free-rides on the prices delivered by the central marketplace.

**Implications for Market Structures and Networks**

The availability of limit orders and/or market maker bids and offers breaks down for large orders such as a 1,000,000 share buy order. Yet, the full market network is robust enough to accommodate the order. What is going on? The divergent expectations hypothesis can shed light on how the big orders are able to get executed. Let’s look at an example. Assume a broker works the order as agent, finds the contra-side, and brings the natural buyer and the natural seller together. What is motivating the buyer? What is motivating the seller? Most likely, neither the big buyer nor the big seller is a noise trader or a liquidity trader. Here is a simpler, more straightforward answer: the buyer has a relatively high valuation and the seller has a relatively low valuation.

The contra-side traders may find one another in an alternative trading system such as Pipeline or Liquidnet. Large orders traded on these systems are generally not from liquidity traders and presumably they are never from noise traders. In addition, by focusing their system on the needs of buy-side customers, Liquidnet does not include sell-side market makers (who could be seeking to trade to rebalance their inventories in specific stocks). Accordingly, the classical trichotomy that we previously discussed breaks down in this market environment. This leads us to replace the homogeneous expectations assumption with the divergent expectations assumption. That is, we presume that the large trade in a single stock results from a natural buyer with a relatively
high valuation meeting a natural seller with a relatively low valuation, and the two of
them implicitly agreeing to disagree.

In keeping with reality, divergent expectations imply a more diverse set of traders
than does the classical framework. There are informed buyers and informed sellers, as
well as liquidity and noise traders. The informed buyers and informed sellers could be
retail-sized investors or institutional-sized investors. Across all of them, expectations can
be divergent, and the more diverse reasons for trading most likely lead to better, more
robust markets. But order size disparity can have consequences.

Consider a large institutional buy order being sliced and diced within a purely
electronic order driven market. The simple adaptive valuations framework of the
previous section counts the number of bulls and bears. What if there is one big bull
executing orders hundreds of times? The one big bull often cannot be distinguished from
hundreds of little bulls each buying just once. The sequential trades triggered by the
multiple tranches of the big bull’s order result in higher prices that can either: (1) be
justified, (2) lead to a bubble, or (3) be ephemeral and lead to increased volatility. A
market network that offers institutions more options than a pure electronic order-driven
market to meet each other can provide the institutional bull with a benefit by giving less
weight to this investor in determining prices.

Off-Book Trading and the No Trade Theorem Revisited

Price discovery and quantity discovery have traditionally been thought of as
occurring within a single trading platform. However, they are two distinct functions, and
they can bifurcate. Over 200 years ago, with reference to a little assembly line in a tiny
pin factory, Adam Smith noted that economic gains can be reaped from a division of
labor (i.e., from individual workers being assigned different specific tasks). An ATS
operating side-by-side with a major market center is a classic replication of Adam
Smith’s pin factory, albeit on a vastly larger scale. Namely, institutional traders can go to
an alternative, less transparent site, meet each other, and trade at lower cost. The
alternative site provides quantity discovery while free riding on the price discovery
services provided by the security exchange’s main trading platform.
In the U.S., quantity discovery is increasingly taking place through ATSs such as Liquidnet (which for the third quarter of 2006 averaged 46.5 million shares a day double-counted, with an average trade size of 48 thousand shares) and Pipeline (which, for the same period averaged 20.98 million shares a day double-counted, with an average trade size of 30 thousand shares). The unbundling, or bifurcation, of price and quantity discovery is occurring in both the U.S. and Europe, but it is difficult to go beyond anecdotal evidence to pin this down because it is not easy to determine how much trading takes place “off-book.”

In addition to the U.S. experience, consider, for instance, Deutsche Börse’s main electronic trading platform, Xetra. Currently, there is no trade reporting requirement in Germany and, consequently, the actual amount of off-book trading in German stocks is not known. To gain further insight into the matter, we have obtained data from Clearstream’s Cascade settlement instructions database for 160 actively traded German equity shares for the month of January 2006 and the period July 2003 through June 2005. On-book trading is predominantly “fast market” activity directed to the Xetra electronic trading facility. Off-book trading includes exchange trading that does not go through Xetra or the physical trading floors, as well as upstairs trading in banks, brokerage firms, and ATSs. What is Xetra’s share of total trading volume?

We have computed the euro value of trades that do, and do not, go through the order book for each of the 160 companies that comprise four major German stock indexes: the 30 DAX large cap stocks, the 50 mid cap MDAX stocks, the 50 small cap SDAX stocks, and the 30 TECDA X technology stocks. For all 160 stocks, the equally weighted percentage amount of on-book trading was 54% for January 2006 and an almost identical 56% for the July 2003 through June 2005 period.

---

13 Clearstream is Deutsche Börse’s clearance and settlement organization. We are grateful to Robert Urtheil and Miroslav Budimir for making these data available to us. Deutsche Börse has also provided us with guidance on how to construct a proxy for off-order book activity based on the Clearstream settlement instructions for the same period. We use this to estimate the relative amount of on-book trading; the estimate is best viewed as a “best guess effort.” See Davis, Pagano, and Schwartz (2006) for more details on the German data and the adjustment methodology used in the analysis.

14 The preponderance of trading goes through the electronic platform, but the German order book data also include trades executed on Germany’s physical trading floors.
For the July 2003 – June 2005 period, the euro trading volume that occurred on-book was relatively small (41%) for the 30 large cap DAX stocks, roughly half (48%) for the mid-cap stocks, a clear majority (65%) for the smaller cap stocks, and yet more substantial for the smaller TECDAX stocks (68%). These averages indicate that larger cap stocks trade more heavily off the book, and that small cap stocks trade more heavily on the book. The finding is intuitively reasonable. Large institutional investors (who are typically more active in the larger cap issues) are less likely to direct their block orders to the book. Additionally, retail orders for these larger cap stocks are more likely to be netted against each other and internalized by German banks and brokers.

One might expect more free-riding when institutional investors, retail clients, and their brokers have more confidence in the exchange-discovered prices. The total order flow for the big cap stocks is larger and this translates into more liquid markets and sharper exchange-delivered price discovery. Hence, more free riding for these stocks may naturally ensue. And so, in the classical tradition of Adam Smith, a “division of labor” appears to be operating in the German market in general, and within the large cap sector of the market in particular. Price discovery is taking place on-book and, to an appreciable extent, quantity discovery is occurring off-book.

What does this imply about the informed, liquidity, and noise trader trichotomy, and the no trade theorem? The prevalence of off-book trading in a major, technologically advanced market conveys an important message to students of the equity markets concerning participants’ motives for trading because participants with different trading motives cluster differently in the standard on-exchange and off-exchange environments. A more diverse group meets at an exchange: small retail and large institutional customers, liquidity traders, information traders, and momentum players. Participants who trade away from the exchange do so for specific reasons that can make the off-exchange composition of the individual liquidity pools less diverse. Retail orders can be executed upstairs by a bank or brokerage house that nets customer orders; the netting process predominantly brings small liquidity traders together. Large institutional orders

---

15 In light of the small number of observations in our sample, we did not attempt to test the statistical significance of the across sub-sample differences.
are executed upstairs by a block trader or by an ATS in order to execute speedily and/or to minimize market impact costs.

ATSs such as Pipeline and Liquidnet receive many large single stock orders and basket trading is not prevalent on either facility.\(^{16}\) The sell-side broker/dealer firms are precluded from Liquidnet and, according to company sources, they direct very little proprietary trading to Pipeline. Thus, for the most part, these two ATSs receive orders that are highly unlikely to be placed by either liquidity or noise traders.\(^{17}\) Consequently, according to the no-trade theorem, Liquidnet and Pipeline should not be viable. Yet they are. As noted, average daily trading volumes for Liquidnet and Pipeline are 46.50 million shares and 20.98 million shares, respectively, and their average trade sizes are 48 thousand shares and 30 thousand shares, respectively. A 48,000 share buy order meeting a 45,000 share sell order for an individual stock at the same moment in time in Liquidnet is very likely the result of two large information-motivated institutional participants coming together to trade with each other. This is particularly strong prima facie evidence that much trading is in fact being driven by divergent expectations.

**Implications for Market Structure Regulation**

Market structure regulation has two over-riding objectives. The first is to improve market efficiency, with efficiency assessed in terms of the containment of trading costs. To this end, major attention has been given to reducing commissions, bid-ask spreads, and market impact costs. Our discussion underscores the importance of including two additional efficiency factors – increasing the accuracy of price discovery and the completeness of quantity discovery.

The second regulatory objective is to ensure fairness across participants by establishing a level playing field. One way in which fairness may be achieved is by regulating information release. For instance, Regulation FD (Fair Disclosure) in the United States prohibits a company from leaking news to preferred investors.\(^{18}\) The

---

\(^{16}\) Trade prices are negotiated in Liquidnet, and thus a list would have to be negotiated one stock at a time. This is a tedious process and is done only occasionally.

\(^{17}\) The minimum order size in Pipeline for a large cap stock is 100,000 shares.

concept has also been applied to ensure that one class of investors, small retail customers, is not disadvantaged by a larger, more powerful and presumably more sophisticated class, the institutional customers. It has been pointed out that these two classes are not so distinct in that pension and mutual funds themselves represent large numbers of small investors. More to the point, it is not at all clear that large funds have an information edge over serious retail investors. And, on the contrary, when it comes to total trading costs, it is the retail investor who frequently has an advantage.

A potential buyer of 1,000,000 shares and a potential buyer of 100 shares may have the same high valuation. The buyer of 100 shares can reach out to a national exchange or any other component of the market network to consummate his or her transaction. The buyer of 1,000,000 shares will have a far more difficult time executing his or her order. Given the high execution costs incurred by multi-billion dollar pension, mutual and hedge funds, their raison d’être for being large must lie elsewhere, and it does. A fund’s comparative advantage lies in constructing, managing, and monitoring portfolios that meet the fund’s objectives and investor desires. The institutions bring disciplined approaches that balance expected return, risk, and liquidity in light of a fund’s objectives. Regulators need to be sensitive to these basic institutional functions, the required concomitant trading, and the total cost of trading.

Regulators have also considered fairness in terms of the relative economic strengths of clients represented by investment institutions, and of intermediaries who, in handling orders and providing dealer capital, stand between the ultimate contra-sides to a trade. The transfer of wealth between investors and intermediaries is substantial. Let us consider this transfer in the context of participant expectations.

The homogeneous expectations hypothesis implies, on average, a triple win-win-win situation for liquidity traders, informed participants, and intermediaries. As trades are made, liquidity needs are met, informed positions with their anticipated future profitability are put into place, and intermediaries are paid for bringing buyers and sellers together. The divergent expectations hypothesis on the other hand suggests that many trades are between informed participants who interpret news differently. For these trades, the future will reveal which of the contras wins and which loses, while intermediaries, on
average, continue to win.\textsuperscript{19} This divergent expectations view is far more consistent with observed institutional investment performance: big funds, on average, do not, consistently over time, have an informational advantage. What does this imply about regulatory policy?

The past three decades have seen a succession of regulatory initiatives: the Congressional mandate that precluded fixed commissions and called for the development of a National Market System (the 1975 Securities Acts Amendments); the institution of a best execution requirement (the 1975 Amendments); the SEC’s Order Handling Rules (introduced in 1997); the removal of the NYSE’s off-board trading prohibition, Rule 390 (in December 1999); the change from fractional to decimal pricing which lowered the minimum tick size to one penny (in 2001); and a trade-through prohibition (a principal component of the SEC’s Regulation NMS which is slated to go into effect in 2007). Each of these can be viewed as an attempt to lower the explicit costs of trading, principally commissions and effective bid-ask spreads. The public policy debates concerning these issues and their eventual translation into market structure have not, to our knowledge, been explicitly discussed in terms of the homogeneity, or divergence, of participant expectations, although the homogeneity assumption has been a “gray eminence” behind them. Consequently, not much attention has been given to price and quantity discovery. Why should it be if shares are thought to have unique fundamental values?

The complexities of price and quantity discovery can be appreciated only when it is recognized that expectations are not homogeneous. Unfortunately, if price and quantity discovery are not adequately taken into account, optimal trading networks will not emerge and costs will be unduly high, particularly for institutional customers. For example, the current U.S. securities exchanges are well-designed for retail investors but this market structure makes it difficult for institutional investors to place large orders in an efficient manner. Thus, these institutional investors are increasingly choosing to trade

\textsuperscript{19} Moreover, dealers are in one regard “informed” participants with expectations that may differ from those of their customers. Their expectations, however, apply primarily to reading the information in the order flow rather than in the asset’s fundamentals. The dealer’s risk profile and relationship with a client are further factors that are taken into consideration. The dealer typically does not have a strong opinion about a stock but will still deal. When selling shares to a 100,000-share bullish buyer, the intermediary is hoping that he or she will soon be buying 100,000 shares from a bearish seller. Thus, the dealers, as intermediaries, may themselves depend on their own customers having divergent expectations.
on off-exchange systems. This, in turn, exacerbates the bifurcation between price and quantity discovery. Intermediaries too may suffer, although they will continue to find ways to bring buyers and sellers together, albeit at higher cost. Consequently, the regulatory debates should take first principles of market microstructure into account. Only then will we be able to achieve a network that enhances efficiency for the broad market.

**There is More Than One Game in Town**

Characterizing participant expectations might appear academic and esoteric. Indeed, market structure and public regulatory policy are rarely discussed in terms of the homogeneity or divergence of participant expectations. We suggest that they should be. Unfortunately, academic analysis has, for the most part, assumed that participant expectations are homogeneous and, accordingly, that stocks have unique fundamental values. For instance, the standard asymmetric information dealer models are based on the assumption that some “informed” traders receive information before others who are “uninformed,” and that all investors who are identically informed have identical (homogeneous) expectations. In this context, the asymmetry of information can explain dealer bid-ask spreads.\(^{20}\) Consequently, major attention has been given to items such as bid-ask spreads and tick sizes, while the complexities of price discovery and quantity discovery have been practically unrecognized. While homogeneous expectations may be a necessary and reasonable assumption for some theoretical modeling purposes, it does not characterize real world markets. More critically, analyses of real world markets that are based on the homogeneity assumption can yield misleading answers.

The motive of institutions executing big blocks is almost always information. But a market driven only by trading based on news that everyone assesses identically is not viable, as Milgrom and Stokey (1982) demonstrate. We therefore ask, what explains institutions seeking to trade individual securities with each other in large amounts? We suggest the answer is that they do not interpret news identically. Their expectations, in other words, are divergent. While taking market prices into account, some portfolio

\(^{20}\)For an early discussion, see Bagehot (1971).
managers form relatively bullish expectations and wish to buy, while others form relatively bearish expectations and wish to sell.

A buyer can know that a seller is looking to trade because of his or her own valuation, and the seller can know the same about the buyer. Regardless, each is willing to trade with the other because each has sufficient confidence in his or her own divergent expectation. In effect, the participants are implicitly “agreeing to disagree” about share value. We suggest that agreeing to disagree in a divergent expectations environment should replace the homogeneous expectations framework used in many microstructure applications.

Comprehending the role played by divergent expectations goes to the heart of two questions – what drives trading, and why does market structure matter? Trading is driven, not by a simple trichotomy of informed, liquidity, and noise traders, but also by participants who value shares differently, who are influenced by each other’s opinions and actions, and who may revise their own expectations at any time. Market structure matters because price and quantity discovery are complex processes. They are complex because shares do not have unique fundamental values. Precise fundamental values do not exist because investor expectations are not homogeneous. Rather, because of the vast and enormously complex information sets that asset managers deal with, investor expectations are divergent. Ultimately, it is the complexity of fundamental information that drives the array of market structure issues that we face today. Consequently, with regard to both market structure design and market structure regulation, we might benefit from paying more attention to how asset managers behave when, faced with imprecise information, they form different opinions about share values.
References


