TECHNOLOGY ADVISORY COMMITTEE

TECHNOLOGICAL TRADING IN THE MARKETS
JULY 14, 2010

U.S. Commodity Futures Trading Commission
1155 21st Street, N.W.
Washington, DC 20581

U.S. COMMODITY FUTURES TRADING COMMISSION
ENSURING THE INTEGRITY OF THE FUTURES & OPTIONS MARKETS

Commodity Futures Trading Commission
<u>Technology Advisory Committee</u>

July 14, 2010 Meeting

TAB	DESCRIPTION
1.	Agenda for Meeting, Wednesday, July 14, 2010
2.	Press Release: CFTC Announces Members of the CFTC's Technology Advisory Committee
3.	Technology Advisory Committee Members
4.	J. Bates, *Algorithmic Trading and High Frequency Trading: Experiences from the Market and Thoughts on Regulatory Requirements* (July 2010).
5.	B. Boultwood, *HFT and Algorithmic Trading Issues and Regulatory Considerations* (July 2010).
6.	J. Castura, R. Litzenberger, and Richard Gorelick, RGM Advisors, LLC, *Market Efficiency and Microstructure Evolution in U.S. Equity Markets: A High Frequency Perspective* (April 22, 2010).
7.	A. Chaboud, B. Chiquoine, E. Hjalmarsson, and C. Vega, *Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market*, BD. OF GOVERNORS OF THE FED'L RES. SYS. INT'L FINANCE DISCUSSION PAPERS, No. 980 (October 2009).
8.	J. Cvitanić and A. Kirilenko, *High Frequency Trading and Asset Prices* (March 11, 2010).
9.	B. Durkin, *The Impact of Algorithmic and High Frequency Trading on CME Group, Inc. Markets* (July 2010).
10.	Futures Industry Association, *Market Access Risk Management Recommendations* (April 2010).

TAB 1

CFTC's Technology Advisory Committee
1155 21st Street, N.W.
Washington, DC 20581

July 14, 2010
Agenda

11:45 a.m. to 12:00 p.m.	Check-in with CFTC Reception – Lobby Level
12:00 p.m. to 12:50 p.m.	Lunch with CFTC Commissioners and Technology Advisory Committee Members, 9th Floor Conference Room
1:00 p.m. to 5:00 p.m.	**TAC Meeting, Hearing Room**
1:00 p.m. to 1:10 p.m.	Opening Remarks: Commissioner Scott D. O'Malia, Chairman, TAC
1:10 p.m. to 1:30 p.m.	Opening Remarks of CFTC Commissioners: Chairman Gary Gensler, Commissioner Michael Dunn, Commissioner Jill Sommers, and Commissioner Bart Chilton
1:30 p.m. to 1:40 p.m.	Overview of Meeting and Introduction of Presenters
1:40 p.m. to 2:00 p.m.	FIA's Market Access Risk Management Recommendations, Mary Ann Burns, Executive Vice President, Futures Industry Association
2:00 p.m. to 3:00 p.m.	Discussion of FIA's Market Access Risk Management Recommendations
3:00 p.m. to 3:15 p.m.	Break *(Restrooms are located on the Mall Level)*
3:15 p.m. to 3:35 p.m.	A Perspective on High Frequency Trading (HFT) from RGM Advisors, LLC, Richard Gorelick
3:35 p.m. to 4:00 p.m.	*High Frequency Traders and Asset Prices,* Andrei Kirilenko, Senior Financial Economist, CFTC Office of the Chief Economist
4:00 p.m. to 4:45 p.m.	Discussion of HFT presentations, debate on need for HFT best practices, and next steps for the TAC.
4:45 p.m. to 5:00 p.m.	Concluding Remarks of CFTC Commissioners: Chairman Gary Gensler, Commissioner Michael Dunn, Commissioner Jill Sommers, Commissioner Bart Chilton and Commissioner Scott D. O'Malia, Chairman, TAC

TAB 2

U.S. COMMODITY FUTURES TRADING COMMISSION

ENSURING THE INTEGRITY OF THE FUTURES & OPTIONS MARKETS

RELEASE: PR5849-10

July 12, 2010

CFTC Announces Members of the CFTC's Technology Advisory Committee – Committee to Meet on July 14, 2010 to Discuss Best Practices for HFT/ALGO

Washington, DC – The first meeting of the CFTC's Technology Advisory Committee (TAC), titled "Technological Trading in the Markets," will be held on July 14, 2010 at 1:00 p.m., at the CFTC's Washington, DC headquarters' Hearing Room. The meeting will address the topics of algorithmic (Algo) and high frequency trading (HFT). Andrei Kirilenko, a Senior Financial Economist in the CFTC's Office of the Chief Economist, will present his paper, **High Frequency Traders and Asset Prices**; Richard Gorelick of RGM Advisors, LLC, will present a high frequency trading firm's perspective on HFT; and a representative from FIA will present the paper **Market Access Risk Management Recommendations** regarding standards on direct market access.

Commissioner Scott D. O'Malia, Chairman of the TAC, has requested that members come to the first meeting prepared to debate the impacts Algo and HFT have on the market and whether or not best practices and/or regulatory standards related to Algo and HFT should be implemented by the Commission.

"The new TAC includes members with a wide range of technology expertise in the financial markets and will be charged with keeping the TAC abreast of new technological advances. With the depth of knowledge of the various TAC members, this new committee can play a vital role assisting the Commission's efforts to better oversee the evolution of the derivatives markets," stated Commissioner O'Malia.

Since the last meeting of the TAC five years ago, technology has become a vital component of the futures and derivatives markets. The timing of the first reconstituted TAC meeting coincides with discussions regarding the market events on May 6th as well as the major financial reform bill soon to be signed into law.

Commissioner O'Malia has requested that TAC members submit papers on various topics, including those listed under related documents. The TAC will have a two-year term, during which time the TAC will conduct public meetings, receive written recommendations from its members and the public, as well as submit reports to the Commission. The TAC will inform the Commission of technological opportunities and make policy recommendations to support the Commission's mission. Non-TAC members are encouraged to submit papers to the TAC as well to help inform the debate to **TechAdvisory@cftc.gov**.

What:	Meeting of the Technology Advisory Committee
Where:	CFTC Hearing Room, 1155 21st Street, NW, Washington, DC
When:	Wednesday, July 14, 2010 1:00 p.m. (EDT)
Viewing/Listening Information:	
The CFTC has made the following options to access meeting:	
1.	Watch a live broadcast of the meeting via webcast on www.cftc.gov.
2.	Call in to a toll-free telephone line to connect to a live audio feed.
Call in participants should be prepared to provide their first name, last name and affiliation. Conference call information listed below:	
	Domestic Toll-Free: (877) 312-3898 Leader Name: CFTC CFTC Conference ID: 86980287

Staff Contact
Stephen Humenik
(202) 418-5314
Office of Commissioner O'Malia

Last Updated: July 12, 2010

Media Contacts
Scott Schneider
202-418-5174

R. David Gary
202-418-5085

Office of Public Affairs

TAB 3

Commodity Futures Trading Commission
Technology Advisory Committee Members

Dr. John Bates
Senior Vice President, Chief Technology Officer and Head of Corporate Development
Progress Software

Dr. Bates is Senior Vice President, Chief Technology Officer (CTO) and head of Corporate Development for Progress Software. Dr. Bates is recognized as a driving force behind the emergence of complex event processing (CEP) and the commercial use of event processing applications in business solutions, including capital markets trading, risk and compliance, telecommunications, fraud prevention, and smart logistics. Prior to joining Progress Software, Dr. Bates was the co-founder, president and CTO of Apama (acquired by Progress Software in April 2005). Before Apama, Dr. Bates was a tenured academic at Cambridge University, where he directed the research into distributed computing systems.

Brenda Boultwood
Chief Risk Officer
Constellation Energy

Ms. Boultwood is Senior Vice President and Chief Risk Officer for Constellation Energy. She leads risk management activities for Constellation Energy and its businesses, including defining and assessing enterprise-wide business risks and facilitating proactive decision-making to effectively manage the risks associated with each business line.

Prior to joining Constellation Energy, Ms. Boultwood most recently served as global head of strategy, Alternative Investment Services for J.P. Morgan Chase & Company, where she was responsible for developing strategy for the company's Hedge Fund Services, Private Equity Fund Services, Leveraged Loan Services and Global Derivative Services business lines. During her tenure at J.P. Morgan Chase, she also served as global head, strategic risk management for its Treasury Services group and as global business head, Global Derivative Services of its Alternative Investment Services group. Ms. Boultwood joined J.P. Morgan Chase when it acquired Bank One Corporation in 2003. Prior to the merger, she held risk management positions with Bank One Corporation, having served as head, corporate market risk management and head, corporate operational risk management and then advancing to head, global risk management for its Global Treasury Services group.

Ms. Boultwood also worked with PricewaterhouseCoopers as a senior manager in its Financial Risk Management Consulting Practice and was employed with Chemical Bank Corporation as a financial engineering associate. In addition, she spent six years teaching in the University of Maryland's Master of Business Administration program.

Ms. Boultwood graduated with honors from the University of South Carolina with a bachelor's degree in international relations. She also earned a Ph.D. in economics from the City University of New York.

John Breyault
Vice President, Telecommunications and Fraud Public Policy
National Consumers League

Mr. Breyault joined the National Consumers League in September 2008. Mr. Breyault's focus at NCL is on advocating for stronger consumer protections before Congress and federal agencies on issues related to telecommunications, fraud, technology, and other consumer concerns. In addition, Mr. Breyault manages NCL's Fraud Center and coordinates the Alliance Against Fraud coalition. Mr. Breyault is also Research Director for the Telecommunications Research and Action Center (TRAC), a project of NCL. In his role with TRAC, Mr. Breyault advocates on behalf of residential consumers of wireline, wireless, VoIP, and other IP-enabled communications services.

Prior to coming to NCL, Mr. Breyault spent five years as director of research at Amplify Public Affairs, where he helped launch the firm's Web 2.0-based public affairs practice and focused on producing actionable public policy research. Earlier in his career, Mr. Breyault worked at Sprint in its International Carrier Services Division and at the American Center for Polish Culture in Washington, DC.

Mr. Breyault was a member of the FCC's Consumer Advisory Committee from 2005 to 2007 and served on the Board of the Arlington-Alexandria Coalition for the Homeless. He is a graduate of George Mason University, where he received a bachelor's degree in International Relations.

Dr. Peter Carr
Global Head of Market Modeling
MorganStanley

Dr. Carr is a Managing Director at Morgan Stanley in New York. He is also the Executive Director of the Masters in Math Finance program at NYU's Courant Institute. Prior to his current positions, Dr. Carr headed quantitative research groups at Bloomberg LP and at Banc of America Securities. His prior academic positions include 4 years as an adjunct professor at Columbia University and 8 years as a finance professor at Cornell University. Dr. Carr is currently the treasurer of the Bachelier Finance Society and an associate editor for 8 journals related to mathematical finance and derivatives. He is also credited with numerous contributions to quantitative finance including: co-inventing the variance gamma model, inventing static and semi-static hedging of exotic options, and popularizing variance swaps and corridor variance swaps. Dr. Carr received his Ph.D. in Finance from UCLA.

Michael Cosgrove
Managing Director-Head of Commodities & Energy Brokerage, North America
GFI Group

Mr. Cosgrove started his career with Amerex Oil Associates in 1981 as a broker of international crude oil. In 1986, Mr. Cosgrove became a partner and Managing Director of Amerex's European operations. During the next 20 years Mr. Cosgrove expanded Amerex's business globally until; in 2006 it comprised 250 employees serving a broad range of international energy and petroleum markets from offices. In 2006 Mr. Cosgrove joined GFI in connection with GFI's purchase of the Amerex North American businesses. In the following year the Amerex U.K. and European businesses were sold in a management buyout and subsequently acquired by Tullet Prebon. In 2004, Mr. Cosgrove was named in the Energy Risk Hall of Fame. Mr. Cosgrove became Managing Director and Head of Commodities & Energy Brokerage for GFI in July 2008 and is responsible for all of GFI's commodities & energy business in North America, including its Amerex and StarSupply divisions.

Gary A. DeWaal
Senior Managing Director and General Counsel
Newedge USA, LLC

Mr. DeWaal is a Senior Managing Director and the General Counsel of Newedge, as well as a member of Newedge's governing Executive Committee. Newedge was created on 2 January 2008 from the merger of the Fimat and Calyon Financial Groups (Newedge refers to Newedge Group and all its branches and worldwide subsidiaries. Newedge Group is jointly owned by Société Générale and Calyon). Newedge's worldwide Legal, Compliance and Financial Crimes Prevention (including AML) departments report to Mr. Dewaal.

Mr. Dewaal joined Fimat in March 1995 from Brody White & Company where he served, at various times, as President, General Counsel, Head of Operations and Head of Internal Audit since 1986. Previously Mr. Dewaal worked for the U.S. Commodity Futures Trading Commission's Division of Enforcement in NYC, and Mudge Rose Guthrie & Alexander, a Wall Street corporate law firm whose principal partner was once Richard Nixon.

Mr. Dewaal also taught a course entitled "Trading Derivatives: Practice and Law" for most years from 1996 through 2006, and has also recently been an annual guest lecturer for the International Finance and Law program in New York City of the State University of New York at Buffalo School of Law.

Mr. Dewaal graduated in 1980 with JD and MBA degrees from the SUNY Buffalo and in 1976 from the State University of New York at Stony Brook where he received a BA degree in English and economics; was elected to Phi Beta Kappa and Omicron Delta Epsilon (international economics honors society); and co-received awards as the University's top overall graduating senior and junior.

Mr. Dewaal has published numerous articles on futures and securities industry issues, and frequently lectures or appears as a speaker at futures and securities industry conferences or in training sessions for regulators. His most recent articles are "Time to Clean up after the Party" (The Financial Times, October 14, 2008), "America's Financial Regulation Needs an Overhaul (The Financial Times, October 31, 2007), "Chicago's Merger Has to Protect the Users' Interest," (The Financial Times, November 15, 2006), "Streamlining Regulation" (The Washington Times, August 2, 2005) and "America Must Create a Single Financial Regulator (The Financial Times, May 19, 2005). He most recently participated as a panelist or speaker on the following panels: Moderator: "Soup to Nuts: General Futures Overview," Futures Industry Association, Annual Law and Compliance Conference (May 2009); Panelist: "Crisis Management: What Happens When Global Firms Fail," International Regulators Symposium and Training Conference, Commodity Futures Trading Commission (March

2009); Moderator: “How Have Current Market Conditions Impact Opportunities in Asia,” Futures Industry Association, Annual Futures Industry Conference (March 2009); Moderator: “Lessons Learned from the Fall (2008) that Was,” Annual International Regulators Symposium and Training Conference, Commodity Futures Trading Commission (October 2008); Lecturer: “Observations on Kerviel: A Broker’s Perspective,” Australian Securities and Investment Commission (September 2008); Panelist: “Managing Risk in Volatile Markets, Financial Industry Regulatory Authority (September 2008); Moderator: “International Transactions” and panelist “Internal Controls post Kerviel:” Futures Industry Association, Annual Futures Industry Conference (March 2008); Panelist: “Introduction to Futures Regulation,” New York City Bar Association Continuing Legal Education Program (December 2007); “Crisis Response,” Annual International Regulators Symposium and Training Conference, Commodity Futures Trading Commission (October 2007); and Moderator: “Asian Regulatory Issues,” Taiwan Financial Securities Commission (October 2007) and FIA Asia Conference (October 2007).

Mr. Dewaal lives in Brooklyn, New York with his wife, Myrna Chao, and two daughters, Justi and Nyasia. In his spare time, Mr. Dewaal is an avid photographer, and bicycles and cooks for fun.

Donald F. Donahue
President and Chief Executive Officer
The Depository Trust & Clearing Corporation

Mr. Donahue is Chairman and Chief Executive Officer for The Depository Trust & Clearing Corporation and for three of DTCC’s operating subsidiaries, The Depository Trust Company, Fixed Income Clearing Corporation and National Securities Clearing Corporation. He took this position in 2007, following on one year as President and Chief Executive Officer for DTCC, DTC, FICC and NSCC, and three years as Chief Operating Officer for DTCC and as President and Chief Operating Officer for DTC and NSCC.

Mr. Donahue has been with DTCC and its predecessor organizations since 1986. During his time at DTCC Mr. Donahue has held positions in a variety of areas, serving as head of the depository’s Operations Division from 1995 until 1997, as the depository’s Chief Information Officer from 1997 until 2000, and as head of DTCC’s Customer Marketing and Development Division, with responsibility for strategic planning, product development, IT applications development, and technology infrastructure support and telecommunications, from 2000 to 2003.

Prior to joining the depository, Mr. Donahue worked for five years for Barr Brothers & Co., Inc., a broker/dealer specializing in municipal securities. He worked for the Municipal

Securities Rulemaking Board, the self-regulatory organization governing the U. S. municipal securities markets, from 1977 to 1985, first as Assistant Executive Director and then as Deputy Executive Director. From 1985 to 1986 he was President of two affiliated companies that developed and marketed secondary market credit enhancements for municipal securities.

From May 2004 to June 2006, Mr. Donahue served, under an appointment by Secretary John W. Snow of the United States Department of the Treasury, as Sector Coordinator for the U.S. Banking and Finance Sector in connection with the Treasury Department's responsibilities as lead agency for the Sector under Homeland Security Presidential Directive 7. In that capacity Mr. Donahue served as Chairman of the Financial Services Sector Coordinating Council for Critical Infrastructure Protection and Homeland Security, a private sector group that interacts with the Treasury Department and Federal and State regulators on infrastructure protection and homeland security issues. From April 2005 to April 2006, Mr. Donahue also served as Chairman of the Partnership for Critical Infrastructure Security, Inc., an organization of all of the Sector Coordinators appointed under HSPD-7 that works with the U.S. Department of Homeland Security on critical infrastructure protection matters.

Mr. Donahue has participated in a variety of financial services industry committees and task forces. He currently serves on the Board of Directors of the United Way of New York City, and on the Board of Directors of XBRL US, the nonprofit consortium for XML business reporting standards in the U.S. financial markets.

Mr. Donahue has a B.A. degree in History from Columbia University.

Bryan T. Durkin
Chief Operating Officer and Managing Director, Products & Services
CME Group Inc.

Mr. Durkin has served as Chief Operating Officer and Managing Director, Products & Services of CME Group since February 2010. He is responsible for the company's Products & Services, Global Operations, Technology and Enterprise Computing, and Enterprise Solutions Divisions. Previously, he served as Managing Director and Chief Operating Officer since July 2007. He also led the global integrations following CME's merger with the Chicago Board of Trade (CBOT) in 2007 and CME Group's acquisition of the New York Mercantile Exchange (NYMEX) in 2008. He is a member of the COMEX Governors Committee and a Director of the CME Foundation. Before joining CME Group, Durkin served as Executive Vice President and Chief Operating Officer of the Chicago Board of Trade (CBOT). Prior to that role, he was in charge of the CBOT's Office of Investigations and Audits where he oversaw the audits, financial surveillance, trade practice and market surveillance self-regulatory and enforcement divisions for the exchange. His career with both CME Group and CBOT spans more than 25 years. Durkin holds a bachelor's degree in business administration and an MBA from Lewis

University. He has been an adjunct faculty member of Lewis University's MBA program, teaching courses in organizational behavior and management.

Richard B. Gorelick
Chief Executive Officer
RGM Advisors, LLC

Mr. Gorelick is the Chief Executive Officer of RGM Advisors, LLC, an automated trading firm that he co-founded in 2001. The company applies scientific approaches and computing power to automated trading strategies in multiple asset classes around the world. The company is headquartered in Austin, Texas, and, through a subsidiary, maintains a London office.

Prior to founding RGM, Mr. Gorelick was the Chief Strategy Officer of Deja.com, Inc., which he joined in 1999, as the company's General Counsel. Prior to Deja, Mr. Gorelick was a corporate attorney in Coudert Brothers' New York office.

Mr. Gorelick received a B.A. in international relations from the University of Pennsylvania and a J.D. from the Georgetown University Law Center.

Dr. Michael Gorham
Industry Professor of Finance and Director
IIT Center for Financial Markets
Illinois Institute of Technology

IIT Industry Professor Michael Gorham served for more than three decades as a research economist at the Federal Reserve Bank of San Francisco. Additionally, Prof. Gorham served as vice president of product development, commodity marketing, education and international marketing spanning 18 years at the Chicago Mercantile Exchange. He also has academic and research experience at IIT's Center for Law and Financial Markets, has been editor of the Journal of Global Financial Markets and provides consulting services to international exchanges and regulators. Most recently, Prof. Gorham served as the first director of the Division of Market Oversight for the Commodity Futures Trading Commission.

Simon Grensted
Managing Director, Business Development
LCH.Clearnet

Mr. Grensted is Managing Director, Business Development. He joined the company in 1997 to design OTC Clearing services. The first of these was the OTC interest rate SwapClear service. Since going live in 1999 the service has grown to include, as users, the majority of the global interbank market makers.

The team has gone on to develop OTC services in repo and bonds, energy, freight and equity derivatives. More recently, he has been involved in developing clearing services for Nodal Exchange, the new US power market as well as the newly formed Hong Kong Mercantile Exchange (HKMEx).

Prior to joining LCH.Clearnet, Mr. Grensted was Director of IT and Operations at EuroBrokers, one of London's largest brokers in emerging markets, money markets and derivative instruments.

Previously, for eleven years from 1982, Mr. Grensted was Managing Director of a software and systems vendor designing and implementing financial systems for banks, analytical products for information vendors and image processing systems for space and defence. During this time he was responsible for the development of a number of market leading products. He has also worked for Reuters and Datastream during his career in financial systems and product design.

Mr. Grensted has a BSc from University College London in Electrical and Electronic Engineering and has been a visiting lecturer at Leeds and Sheffield universities.

Jill L. Harlan
Corporate Risk Manager
Caterpillar, Inc.

Ms. Harlan is currently the Corporate Risk Manager responsible for managing the risk associated with Caterpillar Inc.'s foreign exchange exposures related to the Machine & Engine business.

Ms. Harlan joined Caterpillar as a Finance Analyst in 1988. She is a graduate of Western Illinois University with a Bachelors of Business degree in Finance. She has had a diverse career with assignments focused on both core treasury operations and in marketing business units. Prior to her current position, she served as a Region Finance Manager for North American Commercial Division, Enterprise Risk Manager in Corporate Auditing, Treasurer for Asia Pacific Division, Finance Services Manager for Caterpillar of Australia and Foreign Exchange Administrator for Caterpillar EAME in Geneva. Other assignments in Corporate Treasury include Human Resources Manager, Risk Administrator, Machine Orders Analyst and a variety of Finance Analyst positions.

Douglas E. Harris
Managing Director
Promontory Financial Group, L.L.C.

Mr. Harris is a Managing Director in the New York office of Promontory Financial Group, L.L.C., a financial services consulting and advisory firm, where he advises clients in regulatory matters involving risk management, compliance, investment products, derivatives, capital markets and complex structured transactions, and on general corporate governance, internal controls, and strategic advisory matters.
.
Formerly, he was the General Counsel and Chief Operating Officer of BrokerTec Futures Exchange, L.L.C. and BrokerTec Clearing Company, L.L.C. Before joining BrokerTec, Mr. Harris was a partner in the Regulatory Risk Services Group at Arthur Andersen LLP, where he advised commercial and investment banks, hedge funds and futures commission

merchants on regulatory, compliance, risk management, derivatives and capital markets issues.

From 1993 to 1996, he held the position of Senior Deputy Comptroller for Capital Markets at the Office of the Comptroller of the Currency. He was responsible for the regulation and supervision of national bank capital markets activities, including trading, dealing and investing in derivatives and emerging markets instruments, as well as the development of risk management policies and guidelines. He also served on the Inter-Agency Task Force on Bank-Related Derivatives Activities and as senior staff member of the President's Working Group on Financial Markets.

Previously, Mr. Harris served as Assistant General Counsel of JPMorgan and General Counsel of JPMorgan Futures, Inc.

Mr. Harris is a Director of the National Futures Association, the self-regulatory organization for the U.S. futures industry, where he is a member of both the Compliance Consultative Committee and the Audit Committee. He is also a member of the Bar Association of the City of New York, the New York State Bar Association (Structured Products and Derivatives Law Committee), the American Bar Association (Committee on the Regulation of Futures and Derivatives), the Law & Compliance Division of the Futures Industry Association, and the Financial Markets Association.

Mr. Harris received his AB from Harvard College and his JD from Harvard Law School.

Christopher K. Hehmeyer
Vice-Chairman of the Board
National Futures Association

As CEO of HTG Capital Partners, Mr. Hehmeyer provides the strategic direction and leadership for HTG Capital Partners. Having starter his career as a runner on the floor of the Chicago Board of Trade in 1978, Mr. Hehmeyer became a full member of the CBOT in 1981 where he was a floor broker, floor trader, member of the board of directors, and chaired, vice chaired or served on over 40 committees at the exchange. Most recently he was the CEO of Penson GHCO and continues to serve as its non-executive chairman.

Mr. Hehmeyer was one of the founding partners of Goldenberg, Hehmeyer & Co and worked as a managing director of the Virginia Trading Corporation beginning in 1981 and prior to establishment of the GHCO partnership in 1984.

In addition to his duties as CEO, Mr. Hehmeyer serves as vice chairman of the board of directors of the National Futures Association and as vice chairman of the board of the Futures Industry Association.

He is a member of the advisory board for the Master of Science of Financial Engineering Program at Kent State University, the Economics Club of Chicago, and the World Presidents' Organization.

Mr. Hehmeyer has served on a variety of local charity boards including chairman of the exchange chapter of Ducks Unlimited and chairman of the annual LaSalle Street Dinner Dance for the Chicago Area Council of the Boy Scouts of America.

Steven A. Joachim
Executive Vice President of Transparency Services
Financial Industry Regulatory Authority (FINRA)

Mr. Joachim is the Executive Vice President of Transparency Services at FINRA. His responsibilities include the Alternative Display Facility, FINRA's listed equity quote and trade reporting vehicle; the Trade Reporting Facilities, FINRA's joint ventures with Exchanges for printing listed equity trades; and the Over The Counter Equity transparency facilities, including OTC Bulletin Board and TRACE, the FINRA facility for reporting corporate bond trades.

Prior to his arrival at FINRA in 2002, Mr. Joachim was the Senior Vice President, Chief Operating Officer, Chief Strategy Officer and General Manager for Plural from 1997 to 2001. Plural was a custom interactive software development and strategy firm and is now owned by Dell Professional Services. In 1983, he began a nearly 15-year stint with Merrill Lynch. During his career at Merrill Lynch he served as head of Institutional Marketing, First Vice President, Business Architect for Capital Markets and Chief Technology Officer for Global Equity Markets, Director, Floor Brokerage Services and Business Manager, Global Equity Trading. Throughout his career at Merrill, he has managed operations in Asia, Europe and the U.S. From 1981 to 1983, Mr. Joachim worked for Bankers Trust Company as Vice President, Area Consultant for Lending and Money Transfer Operations. He also served as a Managing Consultant with Cresap McCormick and Paget, Inc.

Mr. Joachim is the current Chairman of the International Forum for Investor Education and has served as a member of the Philadelphia Stock Exchange Board of Governors, Board of Directors for Merrill Lynch Specialists, Inc. and Board of Directors for Wilco, Inc. He has also been a member of the Nasdaq Industry Advisory Committee and the American Stock Exchange Upstairs Member Advisory Committee.

Mr. Joachim has an MA in Political Science from Duquesne University in Pittsburgh, PA, and an MS with distinction in Public Management and a BS in Mathematics from Carnegie Mellon University in Pittsburgh, PA.

Peter G. Johnson
Managing Director of Futures & Options
J.P. Morgan

Mr. Johnson is the Managing Director and Global Co-Head of Futures and Options and OTC Clearing for J.P. Morgan Futures Inc. In March 2010, Mr. Johnson was elected treasurer of the Futures Industry Association (FIA). Mr. Johnson also serves as Chairman of the FIA Market Access Working Group.

Dr. Albert S. Kyle
Charles E. Smith Chair Professor of Finance
University of Maryland

Professor Kyle joined the University of Maryland faculty as the Charles E. Smith Chair Professor of Finance at the Robert H. Smith School of Business in August 2006. He earned his B.S. degree in mathematics from Davidson College in 1974, studied philosophy and economics at Oxford University as a Rhodes Scholar from Texas (1974), and completed his Ph.D. in economics at the University of Chicago in 1981. He has been a professor at Princeton University's Woodrow Wilson School (1981-1987), at the University of California's Haas Business School in Berkeley (1987-1992), and at Duke University (1992-2006).

Professor Kyle's research focuses on market microstructure. His research includes topics such as informed speculative trading, market manipulation, price volatility, and the information content of market prices, market liquidity, and contagion. His current research also deals with concepts from industrial organization to model the valuation dynamics of growth stocks and value stocks by applying techniques used to value real options.

His teaching interests include market microstructure, institutional asset management, venture capital and private equity, corporate finance, option pricing, and asset pricing.

He was elected Fellow of the Econometric Society in 2002. He was a board member of the American Finance Association from 2004-2006. He served as a staff member of the Presidential Task Force on Market Mechanisms (Brady Commission), after the stock market crash of 1987. He has been a member of the NASDAQ economic advisory board and the FINRA economic advisory board.

Garry N. O'Connor
Chief Executive Officer
International Derivatives Clearing Group, LLC

Mr. O'Connor is the Chief Executive Officer of IDCG. Prior to joining IDCG, Mr. O'Connor spent seventeen years in the Investment Banking industry, pricing and managing interest rate derivative portfolios. He has held senior positions in Sydney, Tokyo, Hong Kong and New York with Bankers Trust and then Merrill Lynch. During his time at Merrill Lynch, Mr. O'Connor held a number of roles managing interest rate derivatives risk including leading the Australasian interest rate derivatives trading operation out of Sydney, leading the Japanese Yen swaps desk out of Tokyo, and establishing and managing a US Dollar interest rate trading business in Hong Kong.

Most recently he was charged with establishing a North American presence in the European derivatives markets. At Bankers Trust, Mr. O'Connor managed interest rate, foreign exchange, and commodities risk in Auckland and in Sydney. He was also responsible for price making and risk management activities in Australian and New Zealand interest rate derivatives.

As the CEO of IDCG, Mr. O'Connor has testified on Capitol Hill, met with government regulators and spoken at numerous industry forums on the need for central counterparty clearing and the benefits of extending clearing to all markets participants. He previously served as IDCG's Chief Product Officer and remains responsible for designing and implementing IDCG's cleared interest rate derivative products. He has used his experience as an interest rate trader to design IDCG's product specifications to be economically equivalent to over the counter market.

Mr. O'Connor received a BCom (Hons) from Otago University in 1992.

Michael Ricks
Merchandising Manager, North America
Cargill Inc.

From 1999 to present, Mr. Ricks has been the Merchandising Manager, North America for Cargill Incorporated located in Minneapolis, MN. Prior to that, from 1986-1999 Mr. Ricks was at Continental Grain dealing with grain merchandising. Mr. Ricks received his MS Degree, Agricultural Economics, North Dakota State University, Fargo ND in 1986.

Matt Schatzman
Senior Vice President, Energy Marketing
BG Americas & Global LNG

Mr. Schatzman is responsible for marketing BG's global LNG supply, BG's gas, power and NGL products in North America and BG's oil production in Brazil. Prior to joining BG, Schatzman worked at Dynegy where his last position was president and chief executive officer of Dynegy's energy marketing and power generation business.

Mr. Schatzman holds a Bachelor of Arts degree in political science from Yale University.

Thomas Secunda
Chief Technology Officer
Bloomberg LP

Mr. Secunda, one of the founding partners of Bloomberg, has been with the company since its creation in 1982. Since Bloomberg's inception, Mr. Secunda has served as Director of Research and Development, Director of Worldwide Sales and now Director of Financial Products which includes core terminals, trading systems, tradebook, and portfolio, analytics and risk.

Prior to joining Bloomberg, Mr. Secunda was a fixed-income trader at Morgan Stanley from 1981-1982. Before that he worked in systems research at Salomon Brothers.

Mr. Secunda holds both undergraduate and graduate degrees in mathematics from SUNY Binghamton. He is currently on the Board of Directors of Bloomberg, the National Parks Conservation Association both its national board and it's NYC Council, The Nature Conservancy, the Intrepid Museum Foundation and the Westchester County Parks, Recreation and Conservation Board.

Charles A. Vice
President and Chief Operating Officer
IntercontinentalExchange

Mr. Vice is a founding member of IntercontinentalExchange (NYSE: ICE). He has served as Chief Operating Officer since July 2001 and President since October 2005. Mr. Vice works with the executive management team in setting corporate objectives and strategies and has day-to-day responsibility for technology, operations, and product development. Mr. Vice has been a leader in the management and application of information technology in the energy industry for nearly two decades. Prior to the formation of ICE in 2000, Mr. Vice was a Director at Continental Power Exchange (CPEX), an electronic spot market for electric power. Before joining the CPEX startup in 1994, he was a Principal at Energy Management Associates, where he provided consulting services to the electric power and natural gas industries. From 1985 to 1988, Mr. Vice was a Systems Analyst with Electronic Data Systems (General Motors) where he designed and marketed management information systems for auto, airline and financial service industry clients.

Mr. Vice earned a Bachelor of Science degree in Mechanical Engineering from the University of Alabama and a Master of Business Administration from the Owen Graduate School of Management at Vanderbilt University.

Dr. Matthew White
Senior Economist
ISO New England, Inc.

Dr. White is the Senior Economist at ISO New England. His responsibilities include market design and development for ISO New England's $12 billion suite of auction-based electricity markets.

Prior to joining the ISO, Dr. White held faculty appointments at the Stanford University Graduate School of Business, the University of Chicago (Visiting), and the University of Pennsylvania's Wharton School of Business. There he received numerous outstanding teaching awards for his lectures on how markets work. Dr. White's public service includes appointments as a senior staff economist to the U.S Federal Trade Commission and the U.S. Federal Energy Regulatory Commission.

Dr. White was a Faculty Research Fellow for twelve years at the National Bureau of Economic Research, the nation's premier economic think tank. Dr. White's expertise centers on energy markets and electricity market design, including market microstructure, pricing practices, and demand behavior. His research studies appear in leading academic journals, including the *Review of Economic Studies,* the *RAND Journal of Economics,* the *Review of Economics and Statistics,* and the *Brookings Papers on Economic Activity.* He has served as an evaluator and referee for more than 25 peer-reviewed scholarly journals spanning economics, engineering, and political science. He received his Ph.D. in Economics from the University of California, Berkeley in 1995.

Charles F. Whitman
Chief Executive Officer
Infinium Capital Management

Mr. Whitman is a founding partner and Chief Executive Officer of Infinium Capital Management, a proprietary trading firm based in Chicago with additional operations in New York and London.

Mr. Whitman has been involved in the trading industry since 1987 when, at the young age of 17, he was a runner for Produce Grain Inc. From 1987 to 1992 he attended DePaul University where he double majored in Accounting and Finance. While at DePaul, in 1988 he became a clerk in the soybean options pit for Hanley Group. Mr. Whitman then successfully traded full time in the soybean options pit for several years and in 1996 became a partner at Hanley Group. Starting in 1994, while still trading full time and anticipating the rise in electronic trading, Mr. Whitman focused on developing methods of trading away from the exchange floor. From 1999-2001 Mr. Whitman conducted in depth market and business research for what would become Infinium Capital Management. During this research period, Mr. Whitman also taught options seminars for Dr. Van Tharp and mentored and trained several traders that went on to become extremely successful. In June of 2000 Mr. Whitman became a partner at Blink Trading, LLC, which was sold to GETCO in 2002. Since the launch of Infinium 9 years ago, Mr. Whitman has served the firm simultaneously as CEO and Head of Macro Trading. Infinium is widely recognized for its integrity and multi asset class presence, is an established leader in working with exchanges to develop new products and in 2008 was voted the 4th best place to work in Chicago by Crain's Chicago Business.

Mr. Whitman has been designated a "Super Trader" by Dr. Van Tharp who was featured as the premier psychologist in the trading arena in the book Market Wizards. Mr. Whitman authored the foreword in the current edition of Dr. Tharp's book Trade Your Way to Financial Freedom. Charles Whitman is a longtime member of Chicago Mercantile Exchange, Chicago Board of Options Exchange, Kansas City Board of Trade and Minneapolis Grain Exchange. Furthermore, he is a member of the Economic Club of

Chicago, the Cato Institute and the Chairman's Circle of the Chicago Council on Global Affairs. Mr. Whitman was highly influential in the development of a charitable inner-city ministry, GRIP Outreach for Youth, where he also served as Chairman of the Board. He is devoted to helping at-risk children and has mentored many middle school and high school kids through coaching basketball, one of his lifelong passions. Mr. Whitman is a substantial donor and supporter of several charities including Willow Creek Association, Robin Hood Foundation, Christian Heritage Academy, Caris Pregnancy Centers, World Vision and Direct Relief International.

TAB 4

Algorithmic Trading and High Frequency Trading
Experiences from the Market and Thoughts on Regulatory Requirements
Submission to CFTC TAC on Algo and HFT

Dr. John Bates
CTO, Progress Software and Founder, Apama

Background
This document is divided into 2 sections: firstly it reviews the drivers and trends in Algo and HFT; secondly it discusses the topic of regulation with specific regard to the CFTC, Algo and HFT. I've been very fortune, as the Founder of Apama (one of the leading platforms for Algo and HFT, liquidity aggregation, smart order routing, pre-trade risk and market surveillance) to be involved in working for the last 10 years with leading sell-side and buy-side firms, trading venues and, more recently, regulators. I've seen Algo and HFT evolve in many interesting ways and I wanted to try to capture some of the trends, which also motivate my views of the regulatory requirements going forward.

Algorithmic Trading Terminology
The term algorithmic trading is not used consistently in the industry (sometimes it is used generally and sometimes to describe execution-only strategies or broker algorithms – see below). An algorithm is "a sequence of steps to achieve a goal" – and the general case of algorithmic trading is "using a computer to automate a trading strategy". In almost all cases, algorithms encode what traders could do by watching the market and manually placing orders. The difference is that algos don't need a lunch break or a paycheck! It takes tens of milliseconds for a trader's eye to take in information, communicate with the brain, a decision to be made and the brain impulse to go from a trader's brain to operate his/her fingers to trade. In that time algorithms can have made and executed thousands of trading decisions.

There are 2 main ways in which algorithms are used to automate trading: algorithms for execution and algorithms for HFT.

Execution Algorithms
Execution algorithms are used to break down large orders and slice them into the market over a period of time. The goal is to minimize the impact that a large order has in the market and to achieve a benchmarked price. Examples of this include the VWAP (Volume Weighted Average Price) and Market Participation algos. These algorithms use metrics to determine how to slice a large order; for example, VWAP uses the historic volume distribution for a particular symbol over the course of a day and divides the order into slices, proportioned to this distribution.

The typical use of an execution algorithm is the buyside sending an order to be executed algorithmically into a broker. This can be done either by phone or in an automatic way from a buyside Execution Management System (EMS) as a FIX order. The buyside provides all the information, such as instrument, side, quantity and the algorithm to use. An instance of the execution algorithm is then instantiated within the broker environment to trade the order. It is also possible to run these algorithms within the buyside and just send the *child orders* straight to the market through DMA (direct market access). To achieve this some EMS systems have built-in algorithms and some institutions have built their own algorithms using technologies such as Complex Event Processing (CEP (described later)).

High Frequency Trading Algorithms

While execution algorithms are about automating "how to trade" – i.e. how to place orders in the market, HFT algorithms add to this "when to trade" and even sometimes "what to trade". Execution algorithms are about minimizing market impact and trying to ensure a fair price, whereas HFT algorithms are about profit. The "high frequency" refers to being able to keep up with the high frequency streams of data, make decisions based on patterns in that data indicating possible trading opportunities, and automatically place and manage orders in the market to capitalize.

A term commonly associated with HFT is *statistical arbitrage* (or "statarb") – monitoring instruments that are known to be statistically correlated, with the goal of detecting breaks in the correlation - indicating trading opportunities. For example, consider the relationship (called the delta 1:1) between a bond, such as the 10-year govvie on ICAP (Brokertec), and a derivative of it on CBOT. These instruments tend to move together – but if that relationship breaks for a few milliseconds then there is an opportunity to buy one and sell the other at a profit. There are a variety of types of HFT algorithms for statarb, including the following:

- **Pairs trading** - looking for breaks in the correlated relationships between pairs of instruments.

- **Index arbitrage** - monitoring for breaks in the correlated relationships between instruments and the index of its sector, e.g. Ford against the automotive sector, or a stock index future against one or more of its underlying component elements.

- **Basket trading** - in which statarb techniques are applied not with individual instruments but with custom baskets of instruments.

- **Spread trading** - a related form of statarb that is particularly popular in the futures market. In spread trading, trading is based on taking positions, usually one long and one short, on instruments with profitability being determined by the spread (difference) between two. Examples include the purchase of July Corn and the sale of December Corn (intra-market spread), the purchase of February Lean Hogs and the sale of February Live Cattle (inter-market spread), and the purchase of March Kansas City Wheat and the sale of March Chicago Wheat (inter-exchange spread). More complex inter-exchange multi-legged spreads include *crack spreads*: trading the differential between the price of crude oil and petroleum products, *spark spreads*: trading the theoretical gross margin of a gas-fired power plant from selling a unit of electricity, having bought the fuel required to produce this unit of electricity (and including all other costs operation and maintenance, capital and other financial costs) and *crush spreads*: involving the purchase of soybean futures and the sale of soybean oil and soybean meal futures.

In multi-instrument HFT strategies, low latency is very important – in order to see the patterns in the market and execute trades before competitors. This is particularly relevant in placing multiple trades as part of a statarb scenario – a so-called *multi-legged* trade – where each trade is a leg. Firstly it is important to act on the liquidity opportunity seen in the market; thus fast reaction is important before a competitor takes the opportunity. And secondly, it is important not to get "legged out" – where one leg of the strategy executes but other legs find the market has moved and the opportunity is lost. There are of course mitigating actions that can be taken here, either automatically or manually.

HFT algos are typically used in bank proprietary trading groups, hedge funds and proprietary trading firms. Instances of specific trading strategies can be instantiated by providing key parameters - for example: a new pairs trading strategy needs to know the instruments and specific trading thresholds. Once initiated, HFT algos often run with little human intervention. Often traders monitor the status, P&L and other key parameters on real-time dashboards and can intervene when/if they feel it is necessary. In the case of spread trading, specialized tools called *Spreaders* are often used by traders to instantiate and manage spread trading.

The above HFT algo types are a subset of the algos in the market but illustrate many of the principles. A selection of other areas in which high frequency algorithmic techniques are used include the following:

- **Liquidity aggregation and smart order routing** - As market fragmentation has continued, algorithmic techniques have been employed to aggregate liquidity and use smart order routing to send orders to the venues with the best price and liquidity. These techniques (described below in more detail) can be used by HFT algos to operate more effectively in a fragmented environment.

- **Real-time pricing of instruments** - Algorithmic techniques have also been used in the real-time pricing of instruments, such as bonds, options and foreign exchange. Traditional pricing techniques use slower-moving pricing analytics and fundamentals to price instruments. However, now higher frequency algorithmic techniques can enhance these pricing algorithms based on what is happening in the aggregated market (i.e. how can we make money by increasing the spread on liquidity available) and the tier and history of the customer for whom we are publishing the price (i.e. how should be adjust the spread based on how important the customer is). High frequency pricing can thus skew prices and spreads based on the up-to-millisecond view from the market and the tier of the customer.

- **Trading on news** - In the last couple of years, there has been increasing interest from HFT firms in incorporating news into HFT algos. The idea here is that firms can trade automatically on news sentiment before a human trader can react, e.g. economic releases, news of a war, unexpected weather events etc. They can also correlate and respond to patterns, e.g. the way that news impacts price movements. For a number of years a handful of highly innovative firms have been experimenting with news in HFT. Now, however, this interest is growing due to new types of structured high frequency news feeds. News providers, such as Thomson-Reuters and Dow-Jones are including tags in the feeds that enable algos to quickly extract key information, such as data associated with an economic release.

- **Genetic tuning** - Another interesting technique is *genetic tuning* – in which many thousands of permutations of algorithms are run in parallel and fed with real market data but are not necessarily trading live in the market. The algorithms that have the most profitable theoretical P&L can be put into the market to trade live. Over time live algos may become less profitable and can be deactivated. The branches of profitable algorithms can be grown and the less profitable branches killed off. This model of *Darwinian trading* allows self-evolving systems to discover profitable opportunities through evolutionary processes, with some seeding and guidance by human experts. These techniques are still exploratory and are still only used in a few advanced firms.

The Holy Grail of Algo and HFT is the "money machine" – an algorithm that figures out what to trade and the strategy to trade it, and then continuously self-evolves to remain profitable and outwit competitors. While there are many semi-smart algorithms out there, most still require human expertise and oversight. We are not yet at the stage of truly intelligent algorithms, although it is inevitable this is where the market is aiming for and we must be prepared for this.

The Latency War

In all forms of algo trading – but particularly in HFT, minimizing latency is a key factor in success. Specifically, trading groups are concerned with *end-to-end latency* – the total delay from the market data being generated at the trading venue(s), being delivered to an algo, a decision being taken by an algo and the necessary orders being placed and filled in the venue(s). When several firms are competing for the same opportunity, the one with the lowest latency wins. There is a lot more to Algo and HFT than just latency – as described below – but clearly latency is very important. There are several components in the low latency value chain including:

- **Market data** – traditionally firms like Thomson-Reuters were the preferred one-stop-shop way of delivering market data. However, market data intermediaries can add significant latency and firms focused on HFT are interested in connecting directly to the trading venues through their market data APIs (Application Programming Interfaces). Market data firms have responded by creating lower latency versions of their products and new vendors have emerged such as Wombat (acquired by NYSE) and ActivFinancial.

- **Algo & HFT Engine** – the traditional approach in top tier firms was to hire the top talent and hand-build algorithms in-house using a traditional programming language, such as C++. These algorithms would be tuned to minimize latency in response to patterns in market data. However, with the requirement for quicker time-to-market of new algorithms, new technologies, such as Complex Event Processing (CEP) – which combine rapid development with low latency response to complex patterns in market data have become popular.

- **Order execution** – In recent years, many trading venues have adopted the FIX protocol as the standard way to place orders. In order to minimize latency many institutions connect directly to the venues and place and manage order over FIX.

- **Physical connection** – Some firms have become focused on the physics of reducing latency – making the wire connection over which market data and orders are transmitted as short as possible. There are a number of suppliers, such as BT-Radianz, that can provide a dedicated network that is already wired into trading venues around the world.

- **Co-location** – At the extremes of reducing the latency physics is co-location, in which algorithms are actually installed next to or in the facilities of a trading venue. Several hosting companies have built businesses around providing hosting platforms to allow trading firms to install their software in these *co-lo* facilities. The challenge with co-lo comes for firms that run cross-market, cross-asset or cross-border algorithms – which might involve trading with multiple trading venues that are not geographically co-located. Where does one put these algorithms? Usually at a location with fast inter-connect to all the necessary venues.

Rapid Alpha Discovery, Authoring and Customization

In general, customization of algorithms has been a key differentiator to both brokers that offer execution algorithms and to HFT shops. The principle behind this is that if everyone has the same algorithms then there is no competitive advantage. In practice there isn't a tremendous difference between different broker's VWAP algorithms or different prop shop's pairs trading algorithm – but each firm usually has their own "secret sauce" that makes the algorithm slightly different. There are also obscure algo approaches that are unique to individual firms and are closely guarded secrets.

A key statistic to note, from research by the analyst firm Aite Group, is that the average lifespan of an algorithm is 3 months. This indicates the pace of change in the market. In fact, during the highly volatile markets of late 2008, some firms changed their algorithms on a daily basis – to anticipate and respond to daily opportunities.

In addition to the run-time concerns around minimizing latency (described above) there are also equally important concerns around rapid research and development of new algorithms and customization of existing algorithms. The reasons for this are as follows:

- **First-mover advantage** - The markets change all the time and new patterns that offer potential to build algorithms around emerge. It's key to be able to build, test and deploy a new algorithm quickly because competitors may have spotted the same opportunity and be trying to trade on it first.

- **Adapting to change** - Changes in the market can also impact the effectiveness of existing algorithms. For example, a HFT algo that was trading on a phenomenon that only one firm had spotted initially, may not be effective any more because various competitors have spotted the pattern and are mining it more effectively. Thus the original HFT algo might now be ineffective or even loss making. In this case it's important to be able to either detect this quickly and then switch it off or customize the algo to improve it.

- **Reverse engineering** – There is a fear in the market that competitors can watch the pattern of orders from a particular market participant, figure out how their algorithms work (so-called *reverse engineering*) and then create algorithms to out-perform them. An occasional fear that the buy-side firms have is that their brokers' prop desks are reverse engineering and then front-running their orders. Of course this would be a breach of regulation and brokers are careful to ensure it doesn't happen – but nonetheless the fear remains.

The requirement for customization and continuous innovation is why there is not a large market for vendors of shrink-wrapped pre-built algorithms. However, there is a market for techniques to assist in the rapid creation and customization of algorithms. Trading is an intellectual property business and often the differences in algorithms can be the competitive advantage of one firm over another.

The main areas of interest in rapid development and customization of algorithms are as follows:

- **Alpha Discovery** – Looking for new patterns in the market that might be viable to trade on. Commercial tools exist to assist with this process but also many homegrown tools and techniques are used.

- **Algorithm Authoring and Customization** – Turning a discovered pattern into an algorithm that can trade in the market and then being able to evolve that algorithm over time. There are a number of approaches here. A traditional approach of using an army of developers to code a strategy. This has a number of problems, including slow time-to-market, frequently not coming up with the strategy the business wants and creating a *spaghetti code* maintenance nightmare that can only be understood by certain people who then may leave and create a potential hazard in the market. The term *black box* describes an algorithm the workings of which are hidden. In-house build often creates a black box – only understood by a few technical wizards. Becoming popular is the concept of a *white box* algorithm – which is built on a model, the logic of which can be designed by and is clearly visible to the business and can be easily changed. Modeling tools enable a strategy to be laid out in terms of state flow, rules and analytics. Such tools can generate an executable strategy that can be loaded into an algo engine. The model can be easily changed at any point.

- **Backtesting and Simulation** – Backtesting involves using recorded historic data to determine how an algorithm performs under certain market conditions. This can range from a bull market to a bear market and can use test data from days with certain known phenomena, such as a non-farm payrolls release (or other economic releases) or even the *flashcrash*. Simulation involves providing simulated markets for the algorithms to put their orders into. Simulators can be designed to work hand-in-hand with backtesting environments to, for example, simulate the market impact of trades (because of course we are replaying historic information when the trades we're generating didn't really happen). Different firms vary in their approaches to simulation and backtesting. Some use complex backtesting and simulation; some use simulated markets as provided by trading venues before going live; some prefer iterative testing in live markets.

- **Production** – Putting an algorithm into production involves making it live – so that it is receiving real market data, making trading decisions and placing actual orders in the market. Usually algorithms need to be certified, based on a firm's internal certification procedures, before they are put into production. This often involves extensive backtesting and simulation and often some trader *user acceptance testing* with some selected early-adopter traders. Most trading groups feel backtesting and simulation are not a substitute for real usage – and often reality presents scenarios that were not considered in backtesting. This is why real-time pre-trade risk precautions should be built into all algorithmic platforms to provide additional protection against unforeseen circumstances.

- **Analysis and Tuning** – Once an algorithm has been running live, its performance can be analyzed to detect ways in which it can be optimized to be made more profitable, more efficient or respond more intelligently to certain risk scenarios. This continuous analysis may also discover that an HFT algo is no longer profitable enough and should be modified or discontinued.

The Continuing Evolution of Algo and HFT

One can liken Algo and HFT to gold mining. When gold is discovered in a new territory it's often lying around on the surface. When more people hear about the gold, a goldrush ensues and everyone descends upon the territory. Then one has to pan for gold in rivers or dig to find the hidden seams of gold. In Algo and particularly HFT, trading firms are always seeking out new opportunities and trying to mine them before others descend upon them.

Some key drivers in the evolution of Algo and HFT are as follows:

- **Asset Class** – Initially exchange-traded equities and futures markets were the focus. However, as FX and bond markets have become increasingly electronic, open and fragmentation, so Algo and HFT have grown there. More recently, HFT involving energy trading has been becoming more popular. Throughout this evolution, some firms have employed algos that incorporate cross-asset class trading for statistical arbitrage and hedging purposes.

- **Fragmentation** – As new markets have emerged, so algorithmic techniques have evolved to capitalize. Many asset classes have experienced fragmentation and this trend across all asset classes is likely to continue. Algorithmic techniques to manage fragmentation involve *liquidity aggregation* and *smart order routing*. Liquidity aggregators create a "super book" that combines liquidity on a per symbol or currency pair basis. This offers a global ordered view of market depth for each instrument regardless of which trading venue the liquidity is on. For example, the best bid for a Eurodollar future may be on CME, the second best may be ELX. If human traders or algorithms trade, then smart order routing sends the order to the relevant market(s) on which the quote is displayed. Low latency and rapid update are clearly important here to avoid dealing with stale liquidity information.

  FX is becoming of increasing relevance to the futures community as currency futures and futures equivalents are being aggregated in FX liquidity aggregators, combined with FX-specific trading venues and bank liquidity. Also FX is often used as an important component of cross-border futures strategies.

- **Geography** – Algo and HFT started predominantly in the US and UK markets but have spread geographically over time. Firstly the spread was to other major trading centers such as Tokyo, Sydney, Hong Kong, Toronto and across Europe. Then to locations such as Korea, Singapore and one of the hottest new markets –

Brazil, where both futures and equities are now widely traded algorithmically on BMF-Bovespa. At each stage, algos have to be specialized to the characteristics of the local markets. Each new market presents new trading opportunities.

Algo and HFT Platforms and Technologies

Many firms still use in-house development for the custom creation of algos. However, due to the need to create, evolve, backtest and tune algorithms rapidly, as well as keep up with connections to new trading venues, an increasing number of firms are using third-party products to help accelerate their trading lifecycle. Some key technologies include the following:

- **Execution Management Systems** – Front-end trading systems that allow access to broker algorithms as well as access to custom algorithms integrated with the EMS. Leading providers include FlexTrade, Portware and Orc.

- **Complex Event Processing** – A platform specifically designed for complex analysis and response to high frequency data. CEP platforms, such as Progress Apama, incorporate graphical modeling tools that can rapidly capture and customize strategies and a trading engine connected to any combination of cross-asset market data and trading venues. CEP is used widely for algo trading, HFT, liquidity aggregation, smart order routing, pre-trade risk and market surveillance.

- **Tick Databases** – A real-time time-series database designed to capture and store high frequency data for analysis and backtesting. Providers include Thomson-Reuters and KX Systems.

Algo and HFT Safety Net

HFT can scale the capabilities of a trader hundreds or thousands of times. However, this can of course increase trading risk too. To complement high frequency trading, high frequency pre-trade risk capabilities are needed. Many firms embraced this concept some time ago. However, certain groups used to turn off their pre-trade risk management as it "slowed them down" and any potential downside was over-balanced by the potential upside of trading first. That situation has changed after the 2008 market and the flashcrash, and with increased regulator scrutiny.

Two approaches being successfully used to mitigate trading risk are:

- **Real-time Pre-trade Risk Firewall** – It is possible to continuously recalculate risk exposures on positions whilst monitoring trades as they go to market and determining what impact they would have on pre-defined risk limits. In the event of a threshold breach, trades can be blocked from going to market. It is also possible to monitor for erroneous trades, such as "fat finger trades" and block them. This facility is not just useful for trading groups but also for brokers offering sponsored access, to monitor on a per client basis. Using the latest technology platforms, such as CEP, enables pre-trade checks to be performed with minimal latency.

- **Backtesting and Market Simulation** – As introduced above, before putting algos live, it is highly beneficial to test them with a variety of real historical and pre-canned scenarios to see how they would perform if live. This can be done in conjunction with realistic and tunable market simulators.

Real-time Market Monitoring and Surveillance

Several regulators around the world have recognized that real-time market monitoring and surveillance allows more rapid response to potential crises and market abuse – potentially allowing rapid action to prevent or minimize any market impact. The FSA – the UK regulator - was one of the first to speak up on this and specify a system to achieve more real-time monitoring using Complex Event Processing. Now other regulators around the world are looking at similar approaches. Many trading venues have long had real-time surveillance technologies but there is a lack of consistency across the market. Brokers can also benefits from this kind of technology to prevent abuse in their trading operations and ensure their good reputation.

The goal of real-time monitoring is to detect anomalous market movements, e.g. price or volume spikes for a particular symbol on one or more exchanges. This provides an early warning system to potential market problems and enables rapid response.

In the case of real-time market surveillance, the goal is to detect potential market abuse while it is happening. The FSA drew an analogy that "traders are driving Ferraris and regulators are trying to catch them on bicycles". Utilizing the same technology used in HFT for real-time surveillance and monitoring gives regulators "Ferraris as police cars", to be able to keep up with the high frequency markets. The kind of patterns that can be detected include:

- **Insider trading**, e.g. detection of an unusually large trading pattern followed closely by a news event that moves the market.
- **Front running of orders**, e.g. detection of unusual and coincidental orders from a prop desk just prior to an event that moves the market, such as a large client order being placed by the broker in the same firm or some research being published by an analyst in the same.
- **Painting the tape**, i.e. continuously taking the best offer in the market to drive the price up.
- **Fictitious orders** to manipulate the price and try to get algos to respond.

- **Trader collusion**, in which traders cooperate to deliberately inflate instrument volume and price, such as in *wash trading*.

Keeping an audit trail of market data and potential abuse cases is also important. Tick databases can be used here. Surveillance systems also involve researching new abuse patterns, using business analytics platforms (from providers such as SAS).

Will Algo and HFT replace the trader?

The evolution to Algo and HFT is somewhat analogous to the markets moving from open outcry to electronic trading. When the Liffe floor in London went electronic, half of the traders evolved to the new way of doing things and the other half went to drive London cabs. The same is true with Algo and HFT. Traders who are more involved with simple order entry will inevitably be replaced by technology. However, Algo and HFT is an intellectual property business – so those with the right expertise become the creative minds and the high level coordinators of armies of algorithms. Humans are here to stay.

Are Algo and HFT out of the price range of small firms?

Much has been made in the mainstream press of the unfair advantage of HFT compared with techniques available to the ordinary investor. Actually fund managers will use algos on behalf of the ordinary investor. And HFT helps keep markets more efficient and trading more cost effective at the benefit of the ordinary investor.

With regard to setting up an HFT shop, everything described in this document is available to any firm. The question is can they afford it? HFT is like motor racing. There are some firms that compete in *Formula 1* – with huge budgets and the world's top talent; others compete in national championships; others in club racing; all can potentially win their tier and be successful. For $200,000 or less per year a firm can run a small HFT operation. I believe the costs are going to fall as hosted services offering customizable Algo and HFT capabilities emerge. CQG and Ffastfill, both hosted providers of trading platforms to the commodities and futures markets, have already started offering customizable algorithms and spreaders at a lower cost. It is even possible to offer hosted modeling tools that allow totally custom algos to be created and deployed into the *cloud*. In this kind of scenario small trading firms can concentrate on their own IP and don't need to create in-premise IT shops with skilled IT people, hardware, software, dedicated networks etc. This will inevitably continue to evolve with new entrants, lowering the barriers to entry for HFT.

Thoughts on Regulation of Algo and HFT

Market Impact

What are the positive or negative impacts of Algo and HFT on the futures markets and market structure (e.g. liquidity, volatility; impact of fundamentals, commercials or hedgers; other issues)?

There are a number of positive impacts of Algo and HFT – but there are also a number of potential negative impacts. However, all of the negative impacts can be mitigated by a combination of good policing and best practices from regulators, trading venues and market participants.

The positive impacts of Algo and HFT include the following:

- **Minimize market impact of large trades** – As already described, algorithmic trading provides an automated and intelligence way to break down large orders into smaller chunks to minimize their impact on the market, while achieving a benchmarked price. The market statistics illustrate the impact: for all global markets the average order size has fallen, while the number of order has risen significantly.

- **Lower cost of execution** – Execution algorithms put capabilities previously only available to the elite into the hands of the mainstream buy-side. The use of algos rather than more expensive traders and the competition between brokers continues to drive down margins and help the buyside achieve a significantly reduced cost of execution.

- **More efficient markets** – Most emerging statistical arbitrage opportunities will be quickly identified by firms and algorithms created to mine the seams of gold. Thus markets are continuously evolving and becoming more efficient.

- **More open and competitive trading markets** – Contrary to some popular opinion, there is less of a monopoly in the market generally. Although the top tier firms can still hire top talent and are continuously seen as mysterious controllers of the market, the reality is that one or two individuals can set up a firm that can have access to the same kind of technology as the large players. Technologies like CEP, widely available low latency market connectivity and hosting environments enable “Fred and Ed in a shed” to run an advanced quant trading operation.

- **Faster evolving trading venues** – Market fragmentation has caused increased competition for liquidity between trading venues. This is putting pressure on exchange costs. It is also accelerating the level of technological advancement provides by trading venues – for example, lower matching latency, improved order throughput and more value-added services, such as co-lo.

- **Encouraging entrepreneurship** – HFT is the ultimate form of capitalism. It enables intellectual property to be turned into profit (or loss) rapidly, whether within a large firm or as part of a smaller firm.

- **Increasing productivity** – One trader can manage a handful of instruments and can manage a few trading strategies manually. A trader watches the market and responds by entering orders when (s)he instinctively spots patterns in the market. In an algo-enabled world, a single trader can be the initiator and coordinator of hundreds or thousands of instances of algorithms. The trader can see P&L and status for all algorithmic instances on real-time dashboards and can manually intervene when required. In this way, the productivity of a single trader can be scaled hundreds or thousands of times.

- **Increasing US dominance in the global economy** – Many media commentators have portrayed Algo and HFT as some dark, mysterious, unfair and elite practice. Actually none of these is true or fair. Commentators also speak of the danger of derivatives and we should just go back to owning a share or a commodity – because you know where you stood. In fact the capital markets are a major part of the US economy and a key part of its economic leadership in the world – and central to this is the growing area of Algo and HFT. The recent economic downturn was not caused by Algo and HFT but by more fundamental factors. We can take some lessons from the recent *flashcrash* to enhance the safety system of Algo and HFT, given their importance to the economy. We must be careful not to over-regulate and damage this important economic engine.

Possible negative issues that can arise from Algo and HFT include the following:

- **Accelerating and accentuating market movements** – While algorithms didn't cause the flashcrash, it is likely they accelerated and accentuated it. Algos have no emotion; they are looking for pre-programmed opportunities and will ruthlessly execute against them. Some market panic in particular instruments, as is suspected to have happened in the flashcrash, which then trigger stop-losses and a radical market trend downwards might lead to algos shorting those instruments and then, at an appropriate instant, buying them back at a profit. This is of course true for all market movements every day – and flashcrash-style incidents are infrequent. Although there are a number of mitigating measures that regulators, trading venues and trading institutions can take around real-time market monitoring and response (see below).

- **Easier to *game* the market** – With millions of autonomous algorithms looking for opportunities, it is easier to *spoof* the market, for example by sending in anomalous quotes to try to trigger certain behavior in algorithms. It's also easier to carry out potential market abuse, such as *wash trades* or *painting the tape*, because finding that abuse in a high frequency, fragmented world is challenging. Again here, regulators, trading venues and trading institutions can employ real-time surveillance and response to mitigate these risks.

- **Increased risk profile** – As stated above, algorithms can make a trader hundreds or thousands of times more productive. This can also increase the risk profile hundreds or thousands of times. In addition, algorithms are moving very fast and without proper pre-trade risk precautions, critical exposure levels can be quickly exceeded or errors, such as *fat finger trades,* can be quickly accentuated.

- **Algos can go wild** – Different trading firms have different standards of certification for algorithms before putting them live – and in some cases some logic may be incorrect or missing. Also, the phenomenon of the *black swan* means that algorithms may meet scenarios they have never been prepared for. For these reasons, algorithms can go wrong or behave against their intended specification. This can result in incorrect orders being placed into the market and a large potential loss. Worse still, it can result in a stream of spurious orders being placed into the market. There have been a number of such cases covered in the press in the last year. The problem with algorithms is that they are running at very high speed and detecting these problems can be challenging. One way to catch this is that a trader needs to be watching positions and behavior on real-time dashboard. Ideally algorithmic platforms have a "big red button" to pull one or all algorithms from the market. Often traders then prefer to hedge the undesired positions manually. A more effective approach to algos-gone-wild is to have a real-time pre-trade risk firewall capability – that can block incorrect or spurious trades going to market if they fall outside a particular behavior, break policies or exceed particular risk exposures.

- **Potential for market *denial-of-service*-style attacks** – There have been a number of incidences when out of control algorithms have fired streams of orders into the market in quick succession. This can act in the same way as a network "denial of service" attack – in which a network firewall spends all its time rejecting fraudulent packets and thus cannot accept any real data packets. The market can be taken up with handling these orders and thus slowed down significantly.

- **Additional load on trading venues** – Further to the above point, many algorithms adjust their bids and offers in the markets as the market changes – cancelling current orders and replacing them with modified orders. If this increases it will also start to slow down the markets. Many trading venues have considered charging for excessive order cancellations due to the additional load that it puts on.

- **Increased difficulty of policing the market** – Millions of high frequency algorithms combined with market fragmentation, cross-asset trading, dark liquidity and the challenges identifying which clients of member firms are doing what – all combine to make the job of the regulator very challenging. New technologies and techniques, such as CEP-powered real-time surveillance have been shown to help here – but the situation is still complex.

- **Potentially easier for terrorists to manipulate markets** – A homeland security issue is that if errors, panic and wild algorithms can influence market behavior then it may be possible for terrorists to initiate such behaviors. We need to ensure there are precautions in place to initiate circuit breakers consistently in such a circumstance.

- **Popular fear of "big brother"** – The last year has demonstrated that the media and the general populace have taken a negative attitude to banks and also to HFT. There has been a lot of coverage of certain Senators implying that HFT gives firms an unfair advantage. Unfairly, HFT has been linked to the economic downturn. The media has portrayed *big brother* style algorithms taking advantage of the ordinary investor. Clearly HFT needs a Public Relations makeover!

Regulation and Best Practices

Should the Commission adopt regulations and best practices (e.g. trading, oversight, surveillance and risk management) related to Algo and HFT?

The CFTC should not restrict Algo and HFT. Rather they should improve the policing of the markets in the form of market monitoring and surveillance, and encourage best practices around pre-trade risk for market participants.

What should the role of the Commission, exchanges, clearing organizations and the NFA be with regard to any oversight of Algo or HFT?

The CFTC should take on the role of *God's eye* oversight of the market. In other words the CFTC should be empowered to do the following:

- **Real-time visibility** - See in real-time what is happening on all of the markets the CFTC supervise (and potentially ones they don't). This involves connecting to those markets and getting a real-time feed of full market depth and trade information for each market. Using this information, market monitoring and surveillance should be provided and an audit trail recorded.

- **Real-time market monitoring** - Be able to detect patterns that indicate potentially dangerous market movements, such as price or volume spikes in a particular instrument. This involves parallel monitoring of all instruments on all trading venues. This gives the CFTC an early warning system against potential problems and the ability to immediately communication with trading venues.

- **Real-time market surveillance** - Be able to detect patterns that indicate potential market abuse, such as insider trading or market manipulation in real-time. These incidents should be used to create cases, cross-referenced against past cases and possibly acted upon immediately or later. Real-time visualization is required to show CFTC surveillance staff what is happening in the market and where potential abuse hotspots are occurring.

- **Audit Trail and offline investigation** - Be able to record market and trade data as an audit trail and for additional offline analysis and pattern discovery. One use is to research into the causes of flashcrash-like incidents. Another is to analyze and collect additional evidence for a particular investigation. A further use is to discover new patterns of market abuse so they can be added to the database of patterns and can be looked for in real-time.

- **Improved reporting of OTC products** – Although less real-time, it would be beneficial to improve the accuracy and timeliness of reporting of OTC products. This reporting could then be incorporated into surveillance and analysis operations.

- **Inter-regulator visibility** - In addition, it would be ideal to have cooperation and information sharing between other regulators, both within the US and internationally. For example, there are several trading strategies that may look like market abuse until you see that they are part of a cross-asset strategy involving equities and futures. But the SEC and CFTC in isolation might only see a subset of the trades.

The Exchanges and CFTC should agree and ensure each Exchange has implemented the following:

- **Consistent real-time surveillance and monitoring** - A suitable and consistent level of real-time market surveillance and monitoring for each venue. An Exchange should be able to detect unusual market patterns, such as price and volume spikes, for all instruments. They should also be able to detect gaming and potentially abusive trading patterns in real-time and follow-up with the member firms involved.

- **A consistent definition of when to invoke circuit breakers** - During a market incident (such as a flashcrash), while circuit breakers might work for some instruments on some trading venues, liquidity could just transfer to other venues if they don't have a consistent definition of when to invoke the circuit breakers. Also, circuit breakers should ideally work for all instruments at all times of the day. Additionally the circumstances under which circuit breakers are initiated should be frequently reviewed. In the equities markets there have been a number of false positives since the flashcrash, which has disrupted trading – and the CFTC should consider this in the markets they regulate.

In addition, brokers should implement the following:

- **Pre-trade risk firewalls** – It is highly desirable to analyze each trade in real-time before it hits the market and if necessary block the trade if it is dangerous or erroneous. One use is to ensure a trade does not push exposure beyond key thresholds levels. Another use is to ensure that the trade is not erroneous, such as a "fat finger" trade. A further use is to ensure that an algorithm hasn't gone wild, such as getting stuck in an *infinite loop* whilst sending out trading signals. Such risk firewalls are not just useful for internal users but also for clients in a sponsored access model, to ensure on a client-by-client basis that their pre-trade risk is under control. Clearly it is key that this pre-trade analysis does not slow down HFT algos.

- **Internal market surveillance and monitoring** – It should be possible to detect unusual market movements in particular instruments as an early warning system. Also, each institution should monitor its own trading groups and customers to ensure that the regulator could not perceive any of its trading activity as market abuse.

All of the capabilities mentioned above are possible technologically now. They can be achieved without disrupting current market operations and without putting additional load on the market.

Data Availability

What types of data (i.e. raw feeds) are Algos and HFT receiving from exchanges and news organizations versus what is available to other market participants?

It is not that Algo and HFT traders have privileged access to certain market access methods compared to what is available to other market participants; it is rather that they often need as fast access as possible and choose to pay for it. As already described, HFT shops want to get end-to-end latency as low as possible – so they will want low latency market data as well as low latency market access. A number of market data vendors specialize in low latency market data, and some trading platforms offer direct connectivity to venues, e.g. CBOT or ICE. This access may cost more than less high performing access routes.

Similarly for news providers, there are now news offerings from firms such as Thomson-Reuters and Dow-Jones that offer tagged news for use in HFT. Example uses for news within algos include correlating news announcements with futures movements, or trading on news before the markets, e.g. modifying positions on news of a war, economic event or weather event. Again, these feeds cost extra and are only useful if you are in the HFT space. In fact trading on news in HFT is still a fairly obscure practice that still only a handful of firms are involved in.

Less high frequency participants may gain access to the market through analysis and trading tools, such as Bloomberg or CQG – which have their own integrated market data and news delivery.

It is "horses for courses" when deciding your needs in terms of market data and news. It is certainly not some secret monopoly by Algo and HFT firms. But there is a cost differential in terms of being able to acquire the highest frequency feeds.

Technological Challenges

What are the technological challenges or limitations to Algo and HFT?

When describing the cutting edge of technology, the phrase "rocket science" is commonly used. Ironically rocket science hasn't evolved much in 60 years, whereas algorithmic trading technology evolves daily. In Algo and HFT the barriers of today, described below, will quickly be breached and new barriers will emerge.

Some of the barriers that are being hit or approached by current generation Algo and HFT technology include the following:

- **Transmissions speed (speed of light)** - For those engaged in pure latency arbitrage, the barriers of physics are providing restrictions in terms of their ability to access and respond to market data. Some firms are drilling holes through walls just to shorten the piece of wire that connects them to certain trading markets. Other firms are turning to custom hardware to assist them in pumping the data in quickly. Einstein proved we cannot exceed the speed of light in transmission speeds but in terms of data throughput, data and processing parallelism offer potential for increasing throughput considerably.

- **Analysis speed** – Further to the previous point, keeping real-time analysis up with the rates of data throughput is also becoming challenging. The pattern analysis within algorithms is more complex than just getting the data in –

because complex analytics are being done rather than just shipping data. The OPRA feed, one leading benchmark of analysis requirements, has exceeded 1 million events per second. Technologies like Complex Event Processing, combined with parallelization techniques, such as grid computing, need to continue to evolve to keep up.

- **Trading venue performance** – While trading venues continue to consider performance enhancements, increased loads can hinder their performance. We have already seen performance hits at major exchanges under high load. If algos are allowed to continue to modify orders without restriction then the load on trading venues will continue to increase, requiring continued technology enhancement.

- **Keeping up with global market fragmentation** – For a global cross-asset trading organization, keeping up with new markets is an expensive, complex and time-consuming endeavor. However, the promise of "mining new seams of gold before other prospectors arrive" is enticing.

- **Finding new trading opportunities** – Continuously evolving trading operations and looking for new trading opportunities is challenging. Complex research tools and human processes to help discover new trading opportunities must continue to evolve.

- **Skills shortage** – In order to compete, many trading strategies are becoming more and more complex. For example, a HFT algo that trades on cross-asset, cross-border aggregated liquidity. The human intellectual property and skills in putting together such scenarios and the integrated technology to support them is in short supply.

- **Cost barriers** – To compete at the highest level in Algo and HFT is costly. However, as already described it is possible to get in the game for lower cost.

Conclusions

Algo and HFT are highly beneficial to the US and global economy. Restricting their usage is dangerous. However, best practices and guidance to trading institutions is needed. Mandated pre-trade risk practices, along with market surveillance and monitoring are needed to protect against the potential of further flash crashes and algos going wild - both of which negatively influence market performance and reputation.

TAB 5

HFT and Algorithmic Trading Issues and Regulatory Considerations

Approach: After establishing a definition of High Frequency Trading (HFT), summarize five key issues and discuss potential regulatory considerations.

Definition: High Frequency Trading[1] - use of special software that works in milliseconds to make trades based on market changes

CFTC regulatory goals: fair, orderly and transparent markets

Other possible regulatory objectives: consumer protection, global harmonization, regulatory information sharing, risk assessment and interconnectivity of industry institutions.

HFT and Algorithmic Trading Issues and Description:

Issue	Description
Perceived inequities a. Trader v trader b. Trader to exchange 1. Co-location or proximity c. Exchange v exchange (cross market linkages such as stock indices, commodity indices)	a. Certain market participants are advantaged by advanced technology and quantitative capabilities b. Methods used by a trader to communicate and execute on an exchange c. Linkages across exchanges
Disparities in trader algorithmic model sophistication	Models that seek to profit from market imbalances
Technology moves quickly. HFT Technology and Practices are mature: 70% of US Stock volume is HFT (Tabb Group LLC); CME CEO says that 36% of CME volume is HFT (and 40% of E-mini volume, a CME product that tracks the S&P500)	Lack of market transparency makes formal analysis of volumes difficult, but estimates of overall HFT volumes at 50% of the market
Exchange Volumes: US Exchanges are incented to increase trading volumes, as revenue is generated based on the number of transactions executed • Liquidity rebates • Initial margin rates • Control of traded product underlying (e.g., price index)	In order to attract additional volumes, exchanges are exploring new ways to keep/acquire new clients. • Giving traders ¼ of $.01 for providing liquidity to the market • Reducing initial margin requirements compared to competing exchanges • Creating new products which allow traders to trade new indices.
Dark pools – private trading venues operated by brokers that don't display prices. Fast growing and draw volumes away from exchanges. For example, ECNs such as Archipeligo	Dark pools were originally set up in order to allow large Institutional investors to transact large blocks of equity shares (10,000-100,000) at a time as OTC transactions. These are typically set up by broker-dealers to allow institutional investors to get fair prices and discounted brokerage fees.

[1] Definition: Flash Trade- uses sophisticated high-speed computer technology to allow traders to view orders from other market participants fractions of a second before others in the marketplace.

HFT Issues and Potential Regulatory Considerations:

Issue	Potential Regulatory Considerations
<u>Perceived inequities:</u> a. Trader v trader b. Trader to exchange i. Co-location c. Exchange v exchange	a. The market will determine survival; Regulation likely unnecessary. b. "Co-Location" regulation—Consideration should be given to registration requirements for co-located traders in potential areas such as physical server locations, bandwidth, and computer speeds. . Also, reporting requirements o enhance transparency c. CFTC should consider regulations under its Core Principles for Contract Markets that would createminimum standards and operating rules for technology, resiliency, fair business practices relating to HFTs Consider fines for risk "events" impacting public confidence or weaknesses leading to systemic risk
<u>Disparities in trader algorithmic model sophistication</u>	Markets will eliminate inefficient participants. Regulation likely unnecessary.
<u>Technology Moves Quickly</u>: HFT Technology and Practices are mature: 70% of US Stock volume is HFT (Tabb Group LLC); CME CEO says that 36% of CME volume is HFT (and 40% of E-mini volume, a CME product that tracks the S&P500)	Regulation likely unnecessary • HFT provides 3 beneficial services: liquidity, narrower bid/ask spreads and an overall more efficient market
<u>Exchange Volumes:</u> US Exchanges are incented to increase trading volumes, as revenue is generated based on the number of transactions executed • Liquidity rebates • Initial margin rates • Control of traded product underlying (e.g., price index)	CFTC should consider regulations under its Core Priniciples for Contract Markets that create minimum ethical and technological standards geared toward ensuring safety for entire financial market system. Examples of compliance standards might include technology principles, resiliency plans and strong business practices. Appropriate penalties for non-compliance also should be considered.
<u>Dark pools</u> – private trading venues operated by brokers that don't display prices. Fast growing and draw volumes away from exchanges. For example, ECNs such as Archipeligo	The market will rid itself of inefficient dark pools. If they survive and attract counterparties, transactions will migrate to the dark pools. If inefficient, they will eventually die out and will no longer be necessary.

Next steps:

1. Agree definitions
2. Discuss regulatory considerations
3. Discuss Exchange (CFTC Core Principle) standards forfair, orderly and transparent participation by HFTs

Appendix of potential Regulatory Considerations relating to HFTs (illustrative):

- Track transactions of large HFT traders with more than $20MM in daily volume
- HFT transaction fee or tax
- Circuit breaker standards – e.g. when a particular stock or contract has dropped more than 10% in five minutes
- SEC
    a. 10 minute "circuit breakers" for certain stocks
    b. Proposed: Prohibit flash orders
    c. Proposed: Public display of orders presented to dark pools
    d. Proposed a joint order tracking program
- Sen Kaufman proposals:
    a. Guidance on acceptable and unacceptable practices – flash orders (orders displayed for less than a second to market center's customers to get best execution on that venue before revealing the trade publicly), co-location, naked (i.e., unfiltered access where brokers lets clients execute unsupervised trades on exchanges and market centers without risk controls using the broker's computer codes and without knowing who is making the trade)
    b. Create transparency around HFT activities
    c. Define rules and regulations around what constitutes market manipulation for HFT
    d. Pre-trade risk checking mandatory for any firm accessing electronic trading markets
    e. Introduce a cancellation fee for when firms are cancelling more than a certain percentage of their orders
- CFTC Rules
    a. Large Trader Reporting, includes futures contracts (manual)
    b. Proposed June 10, 2010 - Designated Contract Markets, Derivatives Transaction Execution Facilities and Exempt Commercial Markets that list significant price discovery contracts (i.e., exchanges) and/or proximity hosting services to market participants
        i. Equal access – requires that co-location and proximity hosting services be available to all qualifies market participants willing to pay for the services
        ii. Fees – ensure that cost is not used as a means of denying access by pricing them out of the market
        iii. Latency transparency – ensures that the longest, shortest and average latencies for each connectivity option are available to the public
        iv. Third party proximity hosting services providers – ensures exchanges can obtain information about market participants, their systems and their transactions from a third-party sufficient to carry out self regulatory obligations

TAB 6

Market Efficiency and Microstructure Evolution in U.S. Equity Markets: A High-Frequency Perspective

Jeff Castura, Robert Litzenberger, Richard Gorelick
RGM Advisors, LLC

April 22, 2010

1 Introduction

The impact of high frequency trading (HFT) on the U.S. equity markets has received considerable attention in the wake of the financial crisis of 2008. It has been asked whether the increase in the amount of automated trading as a percentage of overall trading activity over the past several years has been accompanied by degraded measures of market health such as liquidity, trading costs, volatility, etc. Uninformed answers to these important questions have the potential to influence policy makers toward actions that are not beneficial to the vitality and efficient functioning of financial markets in the U.S.

This work presents some evidence showing that the U.S. equity markets appear to have become more efficient with tighter spreads and greater liquidity over the past several years; a period that has seen a sizable increase in the prevalence of HFT, and a period during which there has been coincident growth in automation and speed on many exchanges. It has been suggested that HFT now accounts for over half of U.S. equity share volume [1]. With such a large presence in the market, it is important to understand if there are any adverse effects caused by such activity. While the existence of a causal relationship is not proven, evidence is presented which suggests that the U.S. markets have improved in several respects as HFT activity has grown.

One measure of efficiency investigated is the bid-ask spread and it is expected that the presence of more participants, algorithmic and otherwise, will drive spreads down due to competition, thereby decreasing costs to other investors. The results presented in this paper confirm the results of many other studies, showing that bid-ask spreads have come down over time for a broad range of stocks.

Another measure of efficiency is liquidity, representing the ability of investors to obtain their desired inventories with minimal price impact. Again, it is expected that more participants implies a greater amount of liquidity in the markets, a benefit to investors. This appears to be the case as this paper confirms the results of other papers demonstrating an increase in available liquidity over time.

It was shown by Samuelson in [2] that if a stock price is efficient, i.e., the price is fairly valued with all public information, then it must follow a martingale process. As a consequence, an efficient price exhibits no serial autocorrelation, either positive (momentum) or negative (mean-reversion). Measurements are made in this paper that test how closely stock prices resemble a random walk, and improvements are seen for all segments.

A variance ratio test was developed by Lo and Mackinlay in [3] which makes use of the fact that in an efficient market, the variance per unit time of the price of a stock should be constant. This allows ratios of variances over different time horizons to be taken and compared with theoretical expectations where, in an efficient market, these tests would show that there is little or no serial autocorrelation in prices. Another advantage of this type of test is that it does not depend on a particular order of serial autocorrelation, only whether any such autocorrelation is present. These tests, a novel contribution of this paper, demonstrate that for all the data-sets investigated, there is an overall improvement in efficiency in prices over time.

The data-sets used in this study are the Russell 1000 components, consisting of 1000 large-cap and mid-cap stocks, and the Russell 2000 components, consisting of 2000 small-cap stocks. The set of components are taken as of Q4 2009, and no attempt is made to correct for survivor bias, though it may be argued that the nature of this study is not sensitive to such effects.

Additionally, each index is partitioned into two sets; NYSE-listed stocks and NASDAQ-listed stocks. For much of the time period studied, NASDAQ-listed stocks traded primarily on automated, electronic exchanges while NYSE-listed stocks have transitioned from being primarily traded manually on the NYSE to being traded on a more competitive, automated group of electronic exchanges. Therefore the data essentially represents four distinct subsets of stocks, at least from an historical context: large-cap stocks largely traded automatically (approximately 200 NASDAQ-listed stocks in the Russell 1000), large-cap stocks largely traded manually (approximately 800 NYSE-listed stocks in the Russell 1000), small-cap stocks largely traded automatically (approximately 1300 NASDAQ-listed stocks in the Russell 2000), and small-cap stocks largely traded manually (approximately 700 NYSE-listed stocks in the Russell 2000). This partition allows comparisons to be made that help more clearly identify the impact of automation and technology advances on the health of the market.

The raw data is sampled at 1 second intervals for each stock during the period Jan 1, 2006 to Dec 31, 2009 inclusive, representing 16 quarters of data. The first 10 minutes and last 10 minutes of each day are omitted to prevent opening and closing activities from influencing the results. Inside values are used across the NASDAQ, NYSE, NYSE ARCA and BATS exchanges. This represents a significant fraction of all shares traded in the U.S. and so is taken to be representative of overall market activity.

With this data-set a series of statistical tests and measurements are run, designed to reflect the health of the market. Spreads, available liquidity, and transient volatility in the form of variance ratio tests are presented here as these are commonly cited metrics of market efficiency and market quality.

2 Bid-Ask Spreads

Spreads are a cost to trading and, all else being equal, smaller spreads are evidence of a better cost structure for investors. Conversely, market makers and other liquidity providers earn profits through the spread. To that extent smaller spreads imply not only smaller revenues for market makers but also that these participants, by quoting smaller spreads, are more competitive; a sign of a healthy market.

Bid-ask spreads are presented as the mean absolute spread of each of the components of the index, where the absolute spread is defined as the best ask price less the best bid price. There are other common ways to present bid-ask spread data including the use of relative spreads. This formulation is meant to more directly reflect transaction costs for investors caused by the bid-ask spread. Market makers and other liquidity providers commonly adjust their quotes based on market volatility in order to compensate for their increased risk of holding inventory [4]. Therefore a volatility adjustment is commonly done to remove the impact of volatility from spreads, typically making it easier to spot trends in spreads over time. Dollar-value weighting is also sometimes used in an effort to better reflect costs of the spread paid by investors. Equal weighting is chosen here because many of the largest and most liquid stocks are pinned at a spread of one penny.

Each of these adjustments will alter the results to some degree though overall trends are expected to remain, and this is confirmed in the appendix which contains some results with these adjustments made. Also available in the appendix are some bid-ask spread results for the NASDAQ-100 index, consisting of many of the largest stocks listed on the NASDAQ.

Figure 1 presents the mean of the absolute spread over time for the Russell 1000 stocks partitioned into its NYSE-listed and NASDAQ-listed components. This is done to try to isolate differences in behavior over the period studied that may be attributable to structural changes on each of these exchanges. Both groups have seen a reduction in spreads over the period investigated, dropping by about 1.5 pennies for the NYSE-listed stocks and about 1 penny for the NASDAQ-listed stocks. By the end of 2009 it appears the the mean spread of the two groups has converged to approximately the same value, something that could not be said previously.

It is known that the rate of adoption of automated trading on NYSE-listed stocks lagged behind that of NASDAQ-listed stocks. As the NYSE moved to an electronic system to catch up technologically with the NASDAQ, and as other electronic venues began taking market share from the NYSE, spreads in the Russell 1000 dropped more dramatically for the NYSE-listed stocks than the NASDAQ-listed stocks. This also suggests a relationship between the entrance of algorithmic trading with a reduction in spreads, something that is noted for

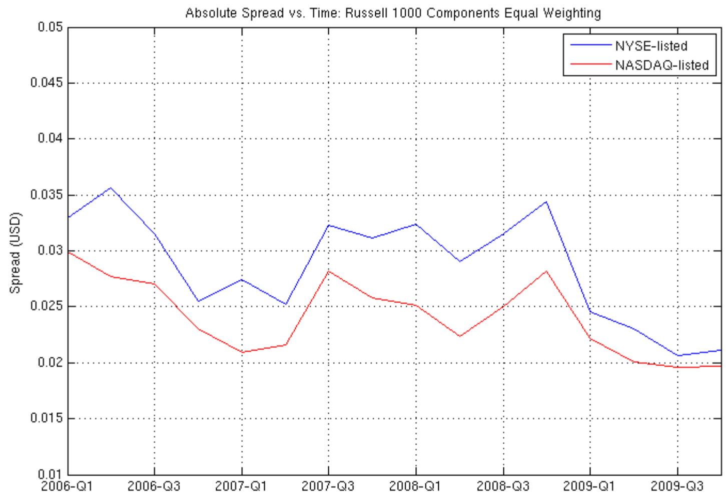


Figure 1: Mean bid-ask spread for Russell 1000

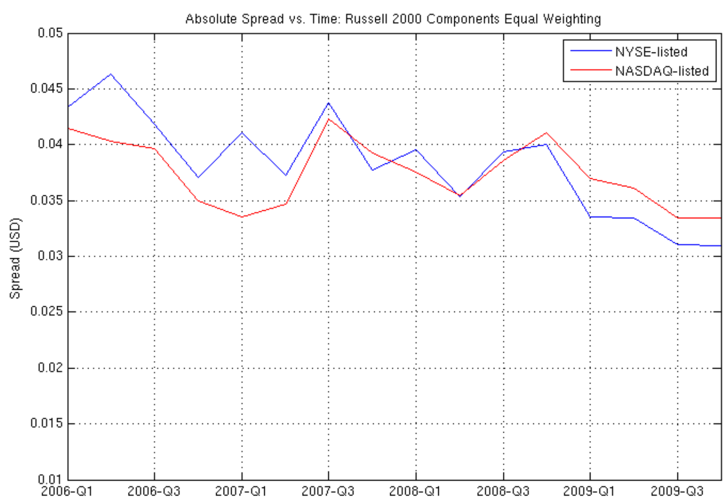


Figure 2: Mean bid-ask spread for Russell 2000

the German DAX in [5].

The same information for the Russell 2000 index is presented in Figure 2. Like the Russell 1000, these stocks have seen a reduction in mean spreads by about a penny, with the NYSE-listed symbols showing a more dramatic reduction than the NASDAQ-listed symbols.

3 Available Liquidity

Liquidity is an important part of a vital market. It is often loosely defined as the ability of participants to trade the amount that they wish at the time they wish. One measure of liquidity is the amount of size offered for sale or for purchase by market makers and other liquidity providers at a given point in time. If more shares are available to be bought and sold at any given time, then market participants have a greater ability to get into or out of positions based on their needs or desires and are less dependent on either waiting for sufficient size to become available or to seek an alternative execution venue.

Available liquidity is measured as the dollar value available to buy or sell at any instant in time at the inside bid and ask, and time averages over an entire quarter are taken. Each stock in an index is weighted by its capitalization reported for the quarter to produce a single capitalization-adjusted available liquidity metric. The motivation for weighting by capitalization is that it more closely reflects the available fraction of a company's total value that can be transacted at any given time which may be more representative of the limitations to investors. Additional available liquidity data is presented in the appendix, including results for the NASDAQ-100.

Figure 3 presents the adjusted available liquidity for the Russell 1000 components partitioned into NYSE-listed and NASDAQ-listed stocks. Between 2006 and the end of 2009, the available liquidity of both groups of stocks increased significantly, by about a factor of two, though all of that gain appears to have taken place in 2009. Similar results are seen for the two groups of stocks in the Russell 2000 which is shown in Figure 4.

It is plausible that the increase in liquidity can be explained, at least in part, by the presence of HFT participants. Since the data used in this work is sampled at a high rate, one can also claim that this liquidity measure is representative of the immediacy that is available to market participants. This immediacy is a type of option that is available to market participants providing them with more flexibility than may otherwise be available.

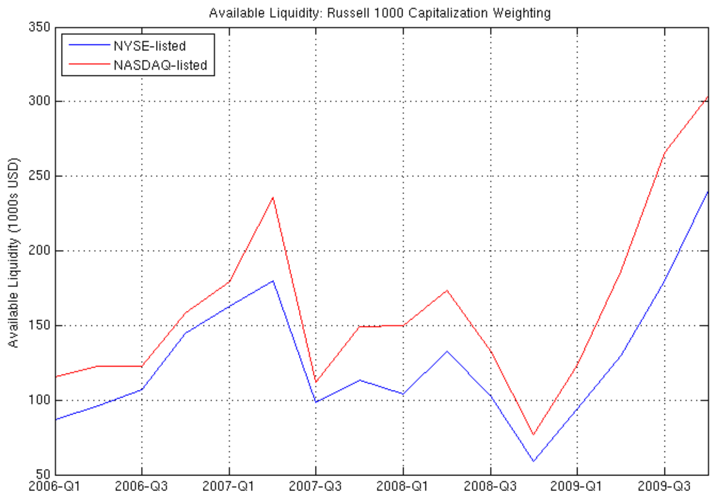


Figure 3: Mean available liquidity for Russell 1000

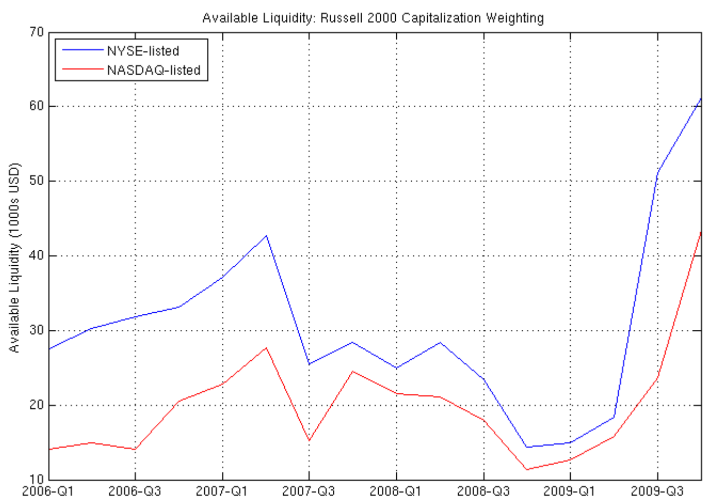


Figure 4: Mean available liquidity for Russell 2000

4 Market Efficiency Tests

There exists a large body of research devoted to tests of market efficiency. In this context, efficiency typically refers to the degree to which the price time-series of a stock resembles a random walk. The theoretical foundation for this was laid out by Samuelson in [2] which proves that a properly anticipated stock price should fluctuate randomly with no serial autocorrelation. Pioneering work in this area in the form of a variance ratio test was presented by Lo and Mackinlay in [3], in which they show with some level of statistical confidence that the NYSE stock-price time-series do not appear similar to a random walk, suggesting inefficiency in the markets. The data used in their paper is sampled daily and ends in 1988, prior to a significant number of structural and regulatory changes that have dramatically changed the nature of U.S. equity markets.

If stock price time-series truly followed random walks, it is expected that the variance ratio computations would have values close to one. A variance ratio's deviation from unity can then be considered to be proportional to the amount of inefficiency present in that stock or index. Values greater than one imply a momentum process, equivalently a positive serial autocorrelation, while values less than one imply mean-reversion, or negative serial autocorrelation.

Subsequent research extended the variance ratio tests in [3] to provide alternative methods to test market efficiency. In particular Chow and Denning in [6] extend the work of [3] to provide a more statistically powerful test procedure and it is this "Chow-Denning" test that is used as a metric of market efficiency in this section. To the best of the authors' knowledge, such tests have not previously been applied to data sampled at a high rate as is done here.

It is important to note that at this sampling rate micro-structural effects are expected to be present. In particular, bid-ask bounce and statistical influences caused by the discrete nature of price values will tend to skew the results toward appearing mean-reverting. These effects are expected at high sampling rates and are expected to decay as the sampling rate is decreased. However, for a given sampling interval, the effect is expected to be roughly constant over time, and thus the interesting aspect of the results is how they have changed over time and whether they have converged toward a value of one. An attempt has been made in the variance calculations to account for the discrete price values and midpoint prices are used rather than last trade prices to minimize the effect of bid-ask bounce. More details are available in the appendix, along with some results based on last trade prices.

Raw variance ratio tests are applied to the Russell 1000 and Russell 2000, partitioned into NYSE-listed and NASDAQ-listed stocks. Three ratios are chosen to be representative of what may be typical HFT holding periods; 10 seconds over 1 second, 60 seconds over 10 seconds, and 600 seconds over 10 seconds.

Figures 5 and 6 show the raw variance ratios of 10 seconds over 1 second for midpoint price data from the Russell 1000 and 2000, respectively. These indexes are partitioned into NYSE-listed and NASDAQ-listed stocks. At this high frequency, it is seen that the Russell 1000, NASDAQ-listed stocks show a high degree of efficiency, and have been relatively efficient throughout the entire period investigated, with some improvement seen over time. As these stocks have largely been traded electronically for the entire period, such results are expected. The NYSE-listed components, by contrast, show a relatively large amount of inefficiency in 2006, but have increased to over 0.95 by 2009 and now appear to be at least as efficient as the NASDAQ-listed stocks.

The Russell 2000 index in Figure 6 shows the same general trends, though the overall efficiency is lower than the Russell 1000. This is to be expected since the smaller-cap stocks of the Russell 2000 do not have the same amount of trading activity as large-cap stocks. The NYSE-listed symbols show a greater degree of improvement in efficiency than the NASDAQ-listed symbols, again coinciding with improvements in automation and increased participation in these stocks by automated trading firms.

The same results are presented for the variance ratios of 1 minute over 10 seconds in Figures 7 and 8 for the Russell 1000 and Russell 2000, respectively. Similar conclusions hold when comparing these results with the 10 seconds over 1 second variance ratios. The degree to which the variance ratio of NYSE-listed stocks in the Russell 1000 has improved of the period studied is dramatic, and has largely converged to be identical to the NASDAQ-listed components of the index. A similar trend is seen with the Russell 2000 components.

A large variance ratio of 10 minutes over 10 seconds is presented to provide a picture of market efficiency over larger time-scales. Figures 9 and 10 show the results for the Russell 1000 and Russell 2000, respectively, and the same general trends seen in the previous plots of variance ratios are present in these figures.

The Chow-Denning method tests the null hypothesis that a price time-series is drawn from a random walk, and produces a single test statistic. This value can be compared to a threshold for a certain significance level. In this study 5% was chosen as the significance level.

The test was applied over each of the 16 quarters, individually to each stock in the data-set with the input to the test being the logarithm of the midpoint price. Sampling was done at 10 minute intervals and 10 second intervals. At 5% significance, if this test were run on 100 truly random time-series, one would expect to see about 5 test outcomes reject the null hypothesis. That is to say, due to the statistical nature of this test, it may produce false positives about 5% of the time.

Results for the 10 minute sampling Chow-Denning tests are presented in Figures 11 and 12 for the Russell 1000 and 2000 data-sets, respectively. These figures show the fraction of stocks in the index that the Chow-Denning test reported as not being drawn from a random walk at a 5%-significance level. Figure 11 shows that at 10 minute sampling, the number of such occurrences has dropped over time and has largely been below 5% since the beginning of 2009, suggesting that there is no statistically significant inefficiencies at this sampling interval that this test detects. The NYSE-listed stocks appear to have a more dramatic improvement, in agreement with the variance ratio results presented above.

Similar results are seen for the Russell 2000 in Figure 12 with a general improvement in efficiency over the time period investigated although it appears

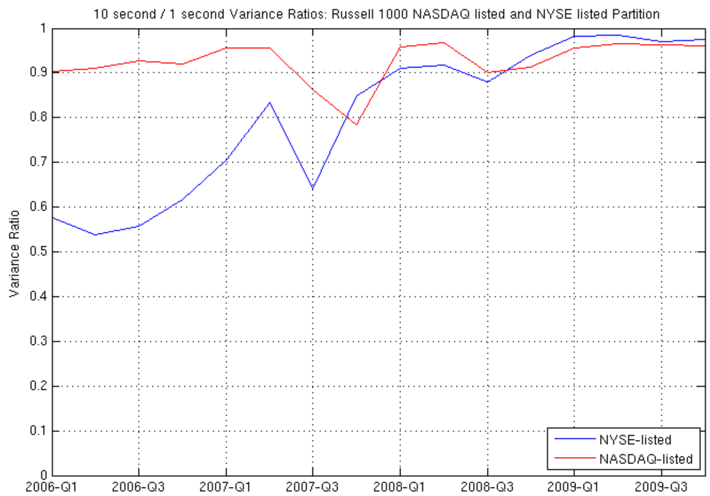


Figure 5: Variance Ratios, Russell 1000, 10 seconds / 1 second

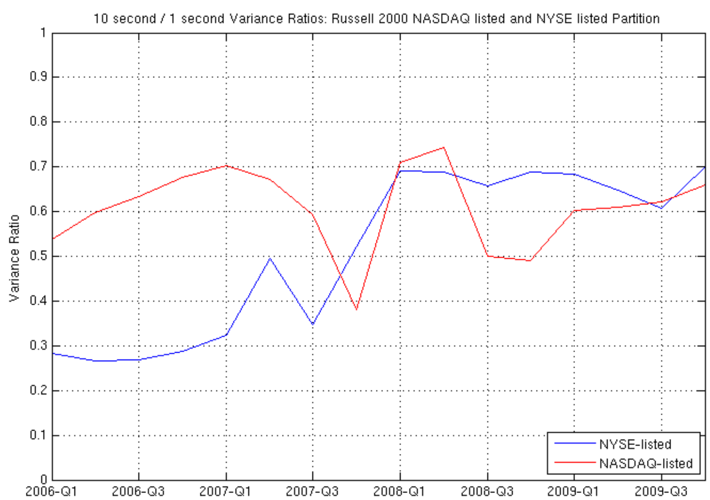


Figure 6: Variance Ratios, Russell 2000, 10 seconds / 1 second

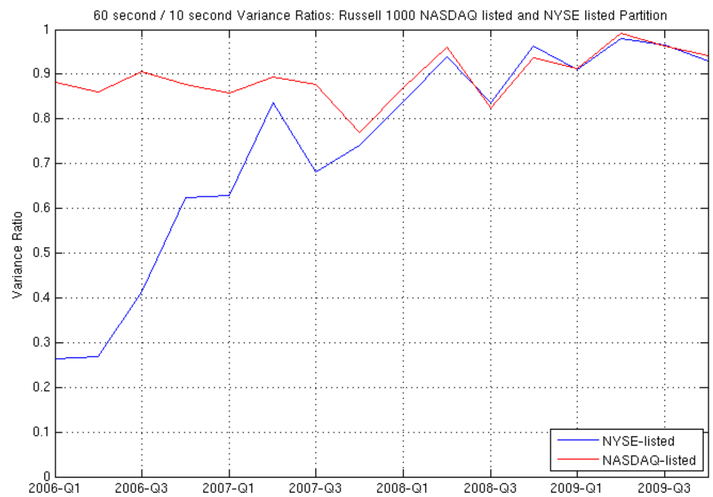


Figure 7: Variance Ratios, Russell 1000, 1 minute / 10 seconds

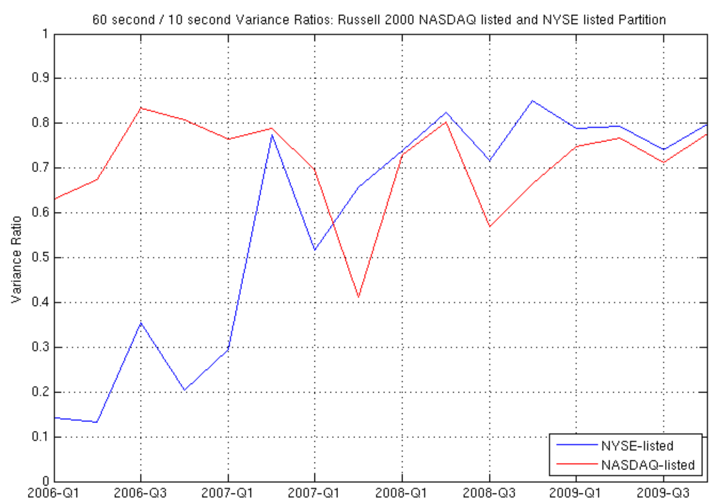


Figure 8: Variance Ratios, Russell 2000, 1 minute / 10 seconds

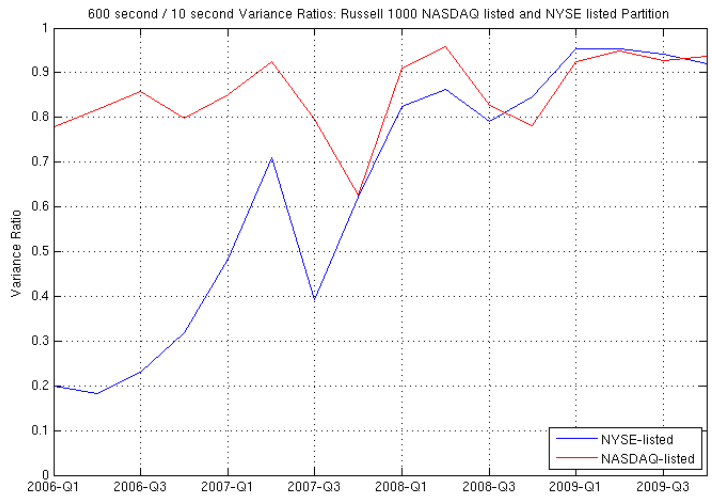


Figure 9: Variance Ratios, Russell 1000, 10 minute / 10 seconds

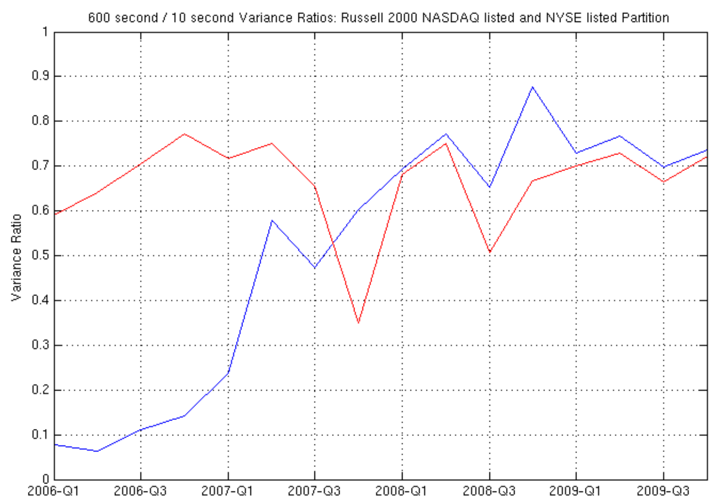


Figure 10: Variance Ratios, Russell 2000, 10 minute / 10 seconds

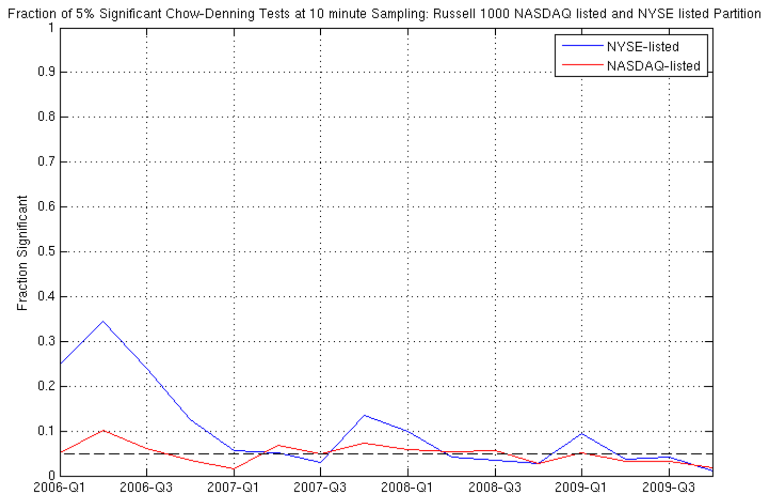


Figure 11: Chow-Denning test results for the Russell 1000, 10 minute sampling

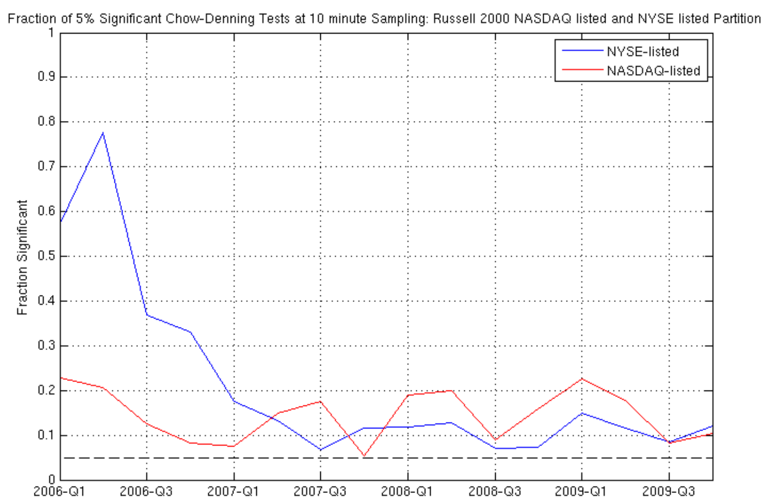


Figure 12: Chow-Denning test results for the Russell 2000, 10 minute sampling

that there remains some degree of inefficiency at this time scale that the Chow-Denning test is detecting. As expected, the large-cap stocks in the Russell 1000 exhibit a smaller number of significant events than the Russell 2000.

A smaller sampling interval of 10 seconds is also used for the Chow-Denning tests, and the results of these computations are presented in Figures 13 and 14 for the Russell 1000 and Russell 2000, respectively. At this sampling rate the impact of microstructural noise is expected to have a more significant impact than at 10 minute sampling. Despite a higher degree of apparent inefficiency, Figure 13 demonstrates that even at such fine sampling, the Russell 1000 appears to have improved over the four years studied, and that the NYSE-listed symbols have shown a more dramatic improvement in that time, largely converging with the NASDAQ-listed symbols. Similar observations are made for the Russell 2000 index in Figure 14.

An alternative interpretation of these results is that of an increase in the speed of mean-reversion over time. As mentioned, mean-reversion is present in this data due in part to micro-structural effects, and as the rate of trading and market activity increases, the impact of such noise on these variance ratio-based tests become less prevalent. Therefore one can conjecture that the decrease in the Chow-Denning test statistics may be as a result of an increased rate of reversion of prices to their mean. This is also an indication of an increasing competitive landscape and increasing efficiency in the market.

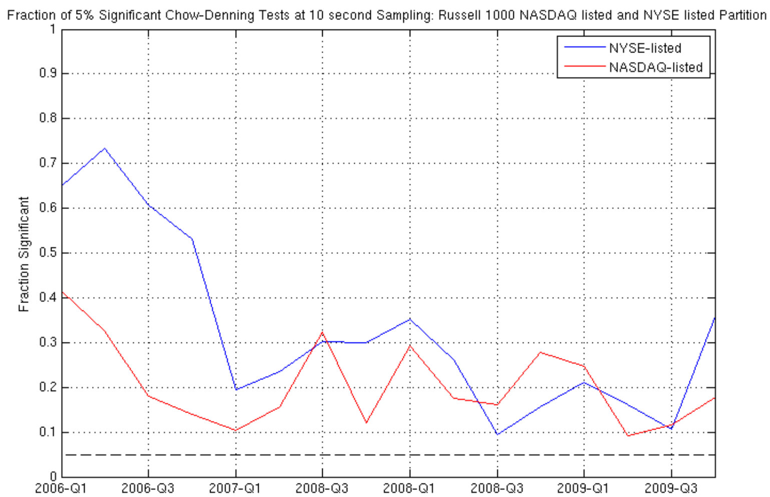


Figure 13: Chow-Denning test results for the Russell 1000, 10 second sampling

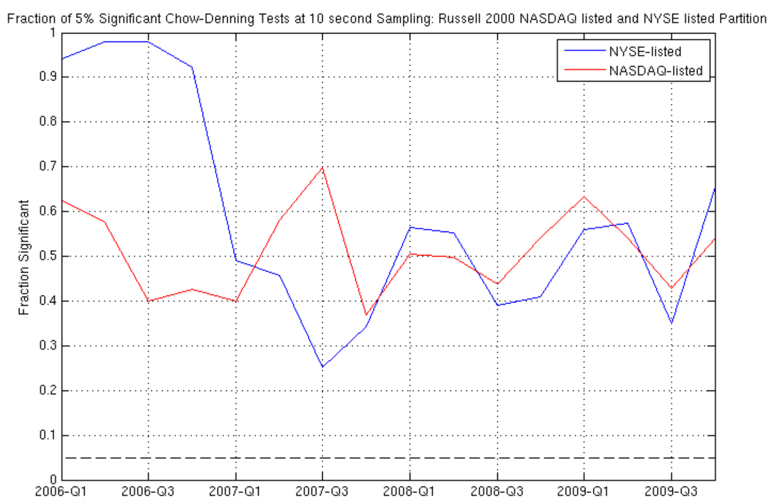


Figure 14: Chow-Denning test results for the Russell 2000, 10 second sampling

5 Summary

The presented data is suggestive that the U.S. equity markets have become more liquid and efficient over the past four years, despite macro-economic shocks. As the ratio of HFT activity to total market activity has grown, there appears to be no evidence that short-term volatility, liquidity or spreads have risen for the bulk of market participants. To the contrary, the evidence presented here suggests a continued improvement in each of these factors, implying a sympathetic relationship between HFT and the health of the overall markets.

The partitioning of data into the Russell 1000 and Russell 2000 shows that there has generally been a larger degree of improvement in efficiency metrics in the Russell 1000. The difference in trends observed between NYSE-listed and NASDAQ-listed stocks also supports the hypothesis that increased automation and the presence of HFT that has come with it has improved the market quality metrics investigated in this paper.

References

[1] The TABB Group, "US equity high frequency trading: Strategies, sizing and market structure", *TABB Group, http://www.tabbgroup.com*

[2] Samuelson, P.A, "Proof that properly anticipated prices fluctuate randomly", *Industrial Management Review*, 1965

[3] Lo, A.W. and Mackinlay A.C., "Stock market prices do not follow random walks: Evidence from a simple specification test", *Review of Financial Studies*, 1988

[4] Glosten, L.R. and Harris, L., "Estimating the components of the bid-ask spread", *Journal of Financial Economics*, 1988

[5] Hendershott, T. and Riordan, R., "Algorithmic trading and information", *http://econpapers.repec.org/paper/netwpaper/0908.htm*, 2009

[6] Chow, K.V. and Denning, K.C., "A simple multiple variance ratio test", *Journal of Econometrics*, 1993

[7] Weisstein, E. W., "Sheppard's Correction", *From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/SheppardsCorrection.html*

6 Appendix

6.1 Bid-Ask Spreads

Absolute spreads are computed as follows. An individual stock i has a spread at time t of $S_i(t) = a_i(t) - b_i(t)$. The spread over a quarter q is defined as

$$\langle S_i(q) \rangle = \frac{\sum_{t \in q} S_i(t)}{\sum_{t \in q} 1}.$$

The spread S_q^N over an index N is the weighted average over all components, where w_i represents the weighting of stock i. The spread is then

$$S_q^N = \frac{\sum_{i \in N} w_i \langle S_i(q) \rangle}{\sum_{i \in N} w_i}.$$

The choice of equal weighting sets all $w_i = 1$. Dollar value weighting is determined by setting the weight for each stock to the total dollar value of all transactions for that stock in the quarter.

Relative spread can be computed in a similar manner, with the relative spread $S_R(t)_i = \frac{a(t)_i - b(t)_i}{p(t)_i}$ replacing the absolute spread above, and where $p_i(t)$ represents price. A common adjustment made to bid-ask spreads is a volatility adjustment [4]. The VIX is used for this purpose and its value relative to the mean of its value over the time period studied is chosen as the deflator. The value of the VIX over the period studied is given in Figure 15.

VIX-adjusted spread data is presented in Figures 16 and 17 showing the Russell 1000 and Russell 2000 relative spreads over time. Similar to the results presented in the main body of this paper, the relative spreads have been stable or falling over time, with a much larger reduction seen when adjusting for the VIX.

For comparison, spread data is also presented for the NASDAQ-100 index. Absolute spreads, both unadjusted and VIX-adjusted are given in Figure 18. The trend for this index is consistent with that seen in the Russell data-sets. Relative spreads are presented in a number of ways in Figure 19 and these adjustments do not change the overall trends presented in the body of the text.

6.2 Available Liquidity

The available liquidity for a stock i at time t is given as

$$L_i(t) = p_i(t) \left(s_i^a(t) + s_i^b(t) \right),$$

where $s_i^a(t)$ and $s_i^b(t)$ are the inside size at the ask and bid, respectively. In a quarter q, the average available liquidity of a stock is

$$\langle L_i(q) \rangle = \frac{\sum_{t \in q} L_i(t)}{\sum_{t \in q} 1}.$$

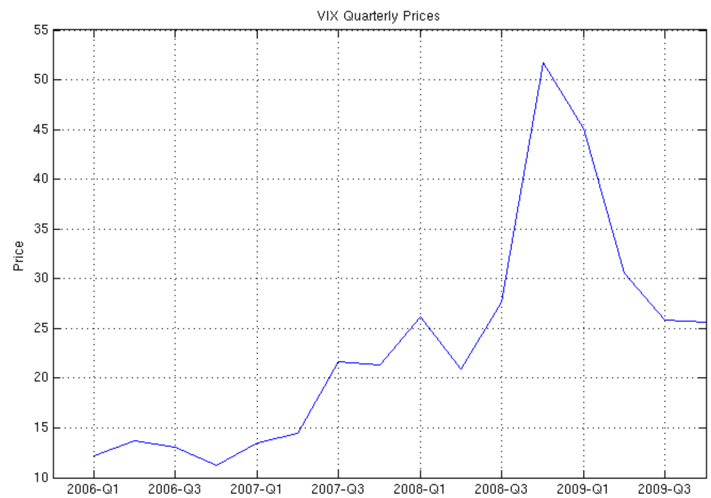


Figure 15: Quarterly VIX prices

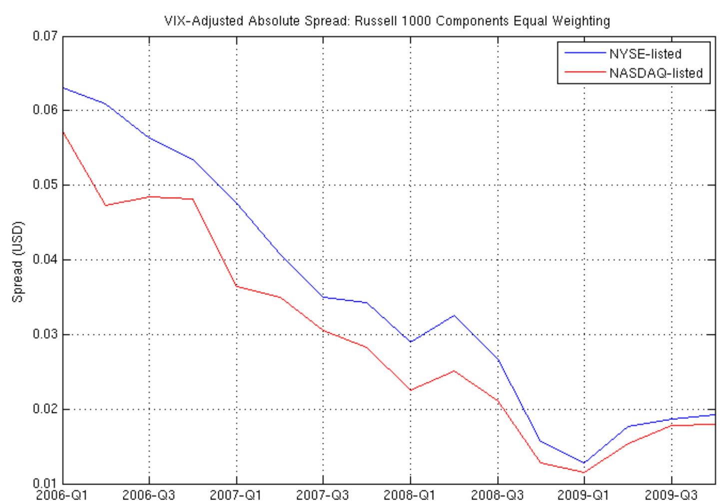


Figure 16: Mean bid-ask spread for Russell 1000, VIX-adjusted

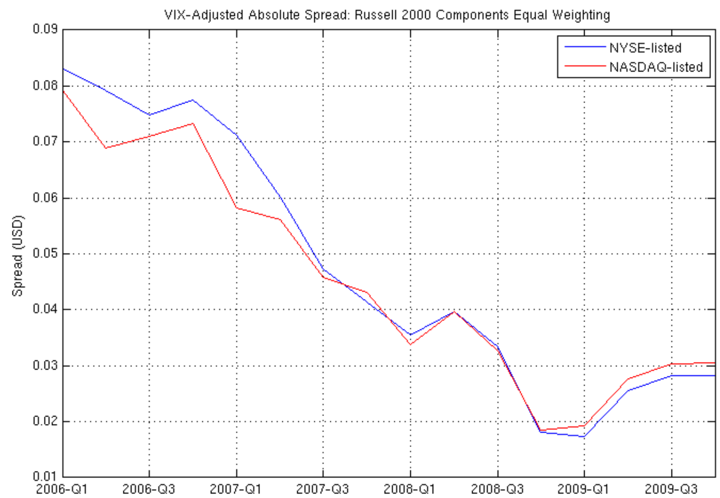


Figure 17: Mean bid-ask spread for Russell 2000, VIX-adjusted

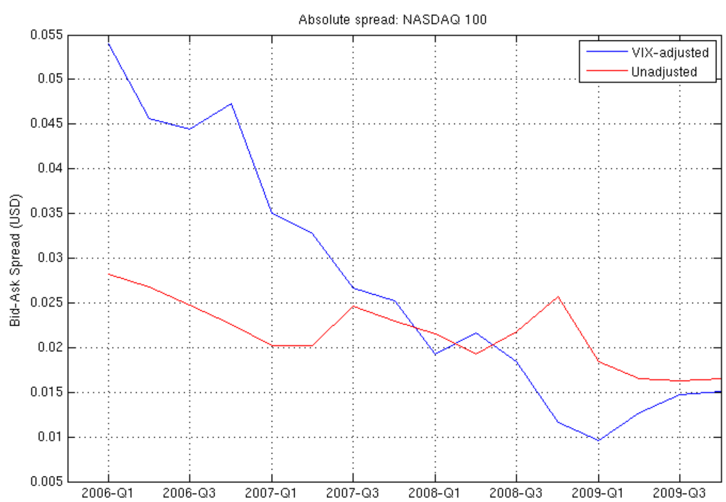


Figure 18: Absolute equal-weighted bid-ask spread for NASDAQ 100

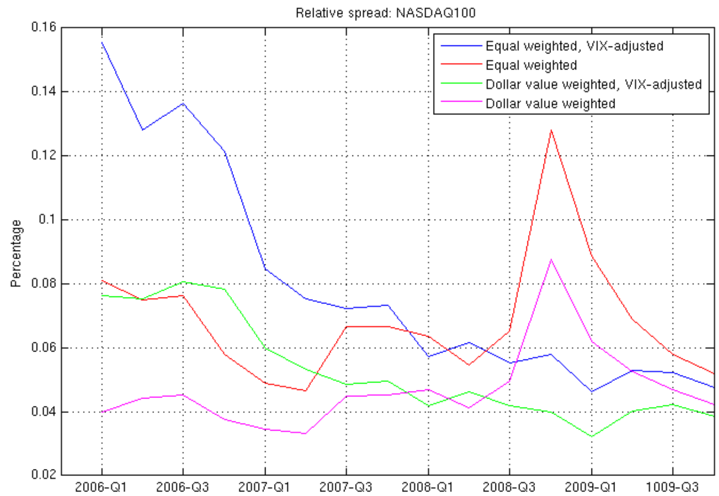


Figure 19: Bid-ask spread for NASDAQ 100

The available liquidity over an index N is the weighted average over all components, such that

$$L_q^N = \frac{\sum_{i \in N} w_i \langle L_i(q) \rangle}{\sum_{i \in N} w_i},$$

where w_i is the weighting for stock i. A common adjustment made is a capitalization adjustment, which is done by setting w_i to the market capitalization of a stock i in quarter q.

The main body of this paper presents results for the Russell 1000 and Russell 2000. For comparison, the available liquidity for the NASDAQ-100 is presented in Figure 20, showing both a capitalization-weighting and an equal-weighting. In both cases, the general trend of increasing available liquidity over the period studied is seen.

6.3 Market Efficiency

The methodology used to compute the variance ratio values follows that presented in [3]. In particular, equations (12a) and (12b) are used. Sheppard's correction [7] is applied to the variance estimates in order to reduce the discrete values of prices (log-midpoint prices) used in the computation.

The raw variance ratio r_i for a stock i with time-ratio D is given by

$$r_i = \frac{v_i^{s_1}}{D v_i^{s_2}},$$

where v^{s_1} is the variance for sampling rate s_1 and v^{s_2} is the variance for sampling rate s_2 and by convention, $\frac{s_1}{s_2} = D > 1$.

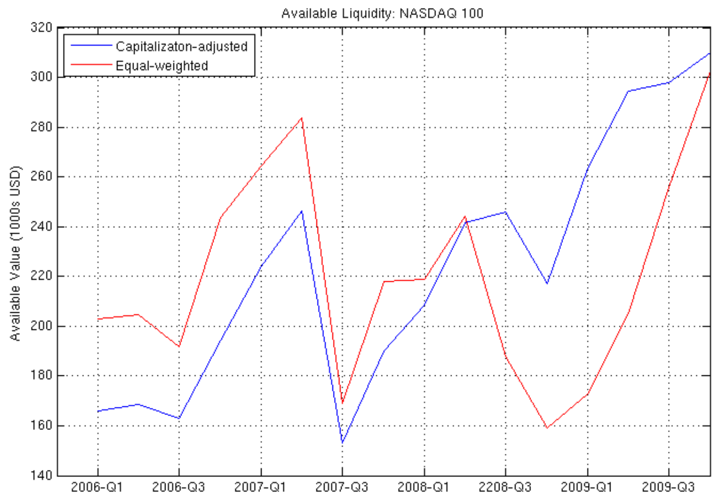


Figure 20: Mean available liquidity for NASDAQ 100

In order to gain a sense of the impact of bid-ask bounce and spreads on variance ratios, Figure 21 presents the raw variance ratios for the NASDAQ 100 using the last traded price and the midpoint price in the same figure. From the left panel, showing a fine sampling rate, it is seen that the impact of the bid-ask bounce on last trade prices results in a smaller variance ratio than when midpoint prices are used. As the sampling rate is decreased to longer time periods, the impact of bid-ask bounce becomes less pronounced. This is demonstrated in the right panel of Figure 21, where the difference between the variance ratios using trade prices and midpoint prices is much smaller.

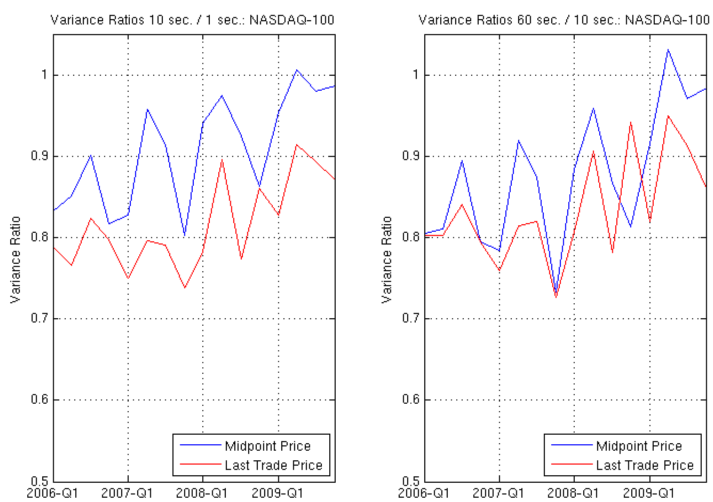


Figure 21: Mean Variance Ratios of Midpoint Prices vs. Trade Prices, NASDAQ 100. Left: 10 seconds / 1 second. Right: 1 minute / 10 seconds

TAB 7

Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 980

October 2009

Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market

Alain Chaboud, Benjamin Chiquoine, Erik Hjalmarsson, Clara Vega

Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market

Alain Chaboud Benjamin Chiquoine Erik Hjalmarsson Clara Vega*

September 29, 2009

Abstract

We study the impact that algorithmic trading, computers directly interfacing at high frequency with trading platforms, has had on price discovery and volatility in the foreign exchange market. Our dataset represents a majority of global interdealer trading in three major currency pairs in 2006 and 2007. Importantly, it contains precise observations of the size and the direction of the computer-generated and human-generated trades each minute. The empirical analysis provides several important insights. First, we find evidence that algorithmic trades tend to be correlated, suggesting that the algorithmic strategies used in the market are not as diverse as those used by non-algorithmic traders. Second, we find that, despite the apparent correlation of algorithmic trades, there is no evident causal relationship between algorithmic trading and increased exchange rate volatility. If anything, the presence of more algorithmic trading is associated with lower volatility. Third, we show that even though some algorithmic traders appear to restrict their activity in the minute following macroeconomic data releases, algorithmic traders increase their provision of liquidity over the hour following each release. Fourth, we find that non-algorithmic order flow accounts for a larger share of the variance in exchange rate returns than does algorithmic order flow. Fifth, we find evidence that supports the recent literature that proposes to depart from the prevalent assumption that liquidity providers in limit order books are passive.

JEL Classification: F3, G12, G14, G15.

Keywords: Algorithmic trading; Volatility; Liquidity provision; Private information.

---

*Chaboud, Hjalmarsson, and Vega are with the Division of International Finance, Federal Reserve Board, Mail Stop 20, Washington, DC 20551, USA. Chiquoine is with the Investment Fund for Foundations, 97 Mount Auburn Street, Cambridge MA 02138, USA. Please address comments to the authors via e-mail at alain.p.chaboud@frb.gov, bchiquoine@tiff.org, erik.hjalmarsson@frb.gov and clara.vega@frb.gov. We are grateful to Terrence Hendershott and Albert Menkveld for their valuable insights, to EBS/ICAP for providing the data, and to Nicholas Klagge and James S. Hebden for their excellent research assistance. We also benefited from the comments of Gordon Bodnar, Charles Jones, Luis Marques, Dagfinn Rime, Alec Schmidt, John Schoen, Noah Stoffman, and of participants in the Spring 2009 Market Microstructure NBER conference, San Francisco AEA 2009 meetings, the SAIS International Economics Seminar, the SITE 2009 conference at Stanford, and the Barcelona EEA 2009 meetings. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

1 Introduction

The use of algorithmic trading, where computer algorithms directly manage the trading process at high frequency, has become common in major financial markets in recent years, beginning in the U.S. equity market more than 15 years ago. There has been widespread interest in understanding the potential impact of algorithmic trading on market dynamics, as some analysts have highlighted the potential for improved liquidity and more efficient price discovery while others have expressed concern that it may be a source of increased volatility and reduced liquidity, particularly in times of market stress. A number of articles and opinion pieces on the topic have recently appeared in the press, with most decrying practices used by some algorithmic traders in the equity market, and there have been calls for regulatory agencies in the United States and Europe to begin investigations.[1] Despite this interest, there has been very little formal empirical research on algorithmic trading, primarily because of a lack of data where algorithmic trades are clearly identified. A notable exception is a recent paper by Hendershott, Jones, and Menkveld (2007), who get around the data constraint by using the flow of electronic messages on the NYSE as a proxy for algorithmic trading. They conclude that algorithmic trading on the NYSE, contrary to the pessimists' concerns, likely causes an improvement in market liquidity.[2] In the foreign exchange market, there has been no formal empirical research on the subject. The adoption of algorithmic trading in the foreign exchange market is a far more recent phenomenon than in the equity market, as the two major interdealer electronic trading platforms only began to allow algorithmic trades a few years ago. Growth in algorithmic trading has been very rapid, however, and a majority of foreign exchange transactions in the interdealer market currently involve at least one algorithmic counterparty.

In algorithmic trading (AT), computers directly interface with trading platforms, placing orders without immediate human intervention. The computers observe market data and possibly other information at very high frequency, and, based on a built-in algorithm, send back trading instructions, often within milliseconds. A variety of algorithms are used: for example, some look for arbitrage opportunities, including small discrepancies in the exchange rates between three currencies; some seek optimal execution of large orders at the minimum cost; and some seek to implement longer-term trading strategies in search of profits. Among the most recent developments in algorithmic trading, some algorithms now automatically read and interpret economic data releases, generating trading orders before economists have begun to read the first line.

[1] See, for instance, "Rewarding Bad Actors," by Paul Krugman, New York Times, August 3, 2009, "High-Frequency Trading Grows, Shrouded in Secrecy," Time, August 5, 2009, and "Don't Set Speed Limits on Trading," by Arthur Levitt Jr., Wall Street Journal, August 18, 2009.

[2] We also note a paper by Hasbrouck (1996) on program trading, where he analyzes 3 months of data where program trades can be separately identified from other trades. He concludes that both types of orders have an approximately equivalent impact on prices. Algorithmic trading is not exactly equivalent to program trading, though it is a close cousin. In principle, a program trade could be generated by a trader's computer and then the trade conducted manually by a human trader. Our definition of AT refers to the direct interaction of a trader's computer with an electronic trading platform, that is the automated placement of a trade order on the platform.

The extreme speed of execution that AT allows and the potential that algorithmic trades may be highly correlated, perhaps as many institutions use similar algorithms, have been cited as reasons for concerns that AT may generate large price swings and market instability. On the other hand, the fact that some algorithms aim for optimal execution at a minimal price impact may be expected to lower volatility. In this paper, we investigate whether algorithmic ("computer") trades and non-algorithmic ("human") trades have different effects on the foreign exchange market. We first ask whether the presence of computer trades causes higher or lower volatility and whether computers increase or reduce liquidity during periods of market stress. We then study the relative importance of human and computer trades in the process of price discovery and re-visit the assumption that liquidity providers are "uninformed."

We formally investigate these issues using a novel dataset consisting of two years (2006 and 2007) of minute-by-minute trading data from EBS in three currency pairs: the euro-dollar, dollar-yen, and euro-yen. The data represent the vast majority of global spot interdealer transactions in these exchange rates. An important feature of the data is that the volume and direction of human and computer trades each minute are explicitly identified, allowing us to measure their respective impacts.

We first show some evidence that computer trades are more highly correlated with each other than human trades, suggesting that the strategies used by computers are not as diverse as those used by humans. But the high correlation of computer trades does not necessarily translate into higher volatility. In fact, we find next that there is no evident causal relationship between AT and increased market volatility. If anything, the presence of more algorithmic trading appears to lead to lower market volatility, although the economic magnitude of the effect is small. In order to account for the potential endogeneity of algorithmic trading with regards to volatility, we instrument for the actual level of algorithmic trading with the installed capacity for algorithmic trading in the EBS system at a given time.

Next, we study the relative provision of market liquidity by computers and humans at the times of the most influential U.S. macroeconomic data release, the nonfarm payroll report. We find that, as a share of total market-making activity, computers tend to pull back slightly at the precise time of the release but then increase their presence in the following hour. This result suggests that computers *do* provide liquidity during periods of market stress.

Finally, we estimate return-order flow dynamics using a structural VAR framework in the tradition of Hasbrouck (1991a). The VAR estimation provides two important insights. First, we find that human order flow accounts for much of the long-run variance in exchange rate returns in the euro-dollar and dollar-yen exchange rate markets, i.e., humans appear to be the "informed" traders in these markets. In contrast, in the euro-yen exchange rate market, computers and humans appear to be equally "informed." In this cross-rate, we believe that computers have a clear advantage over humans in detecting and reacting more quickly to triangular arbitrage opportunities, where the euro-yen price is briefly out of line with prices in the euro-dollar and dollar-yen markets. Second, we find that, on average, computers or humans that trade on a price posted by a computer do not impact prices quite as much as they do when they trade on a price posted by a human. One possible interpretation of this result is that computers tend to place limit orders more strategically than humans do. This empirical evidence supports the literature that proposes to depart from the prevalent assumption that liquidity providers in limit order books are passive.[3]

The paper proceeds as follows. In Section 2 we introduce the EBS exchange rate data, describing the evolution over time of algorithmic trading and the pattern of interaction between human and algorithmic traders. In Section 3 we study the correlation of algorithmic trades. In Section 4 we analyze the relationship between algorithmic trading and exchange rate volatility. In Section 5 we discuss the provision of liquidity by computers and humans at the time of a major data release. In Section 6 we report the results of the high-frequency VAR analysis. We conclude in Section 7. Some robustness results are presented in the Appendix.

2 Data description

Today, two electronic platforms process the vast majority of global interdealer spot trading in the major currency pairs, one offered by Reuters, and one offered by EBS.[4] These platforms, which are both electronic limit order books, have become essential utilities for the foreign exchange market. Importantly, trading in each major currency pair has over time become very highly concentrated on only one of the two systems. Of the most traded currency pairs, the top two, euro-dollar and dollar-yen, trade primarily on EBS, while the third, sterling-dollar, trades primarily on Reuters. As a result, the reference price at any moment for, say, spot euro-dollar, is the current price on the EBS system, and all dealers across the globe base their customer and derivative quotes on that price. EBS controls the network and each of the terminals on which the trading is conducted. Traders can enter trading instructions manually, using an EBS keyboard, or, upon approval by EBS, via a computer directly interfacing with the system. The type of trader (human or computer) behind each trading instruction is recorded by EBS, allowing for our study.[5]

We have access to AT data from EBS from 2003 through 2007. We focus on the sample from 2006 and 2007, because, as we will show, algorithmic trades were a very small portion of total trades in the earlier years.

[3] For example, Chakravarty and Holden (1995), Kumar and Seppi (1994), Kaniel and Liu (2006), and Goettler, Parlour and Rajan (2007) allow informed investors to use both limit and market orders. Bloomfield, O'Hara and Saar (2005) argue that informed traders are natural liquidity providers, and Angel (1994) and Harris (1998) show that informed investors can optimally use limit orders when private information is sufficiently persistent.

[4] EBS has been part of the ICAP group since 2006.

[5] EBS uses the name "automated interface" (AI) to describe trading activity directly generated by a computer, activity we call AT.

In addition to the full 2006-2007 sample, we also consider a sub-sample covering the months of September, October, and November of 2007, when algorithmic trading played an even more important role than earlier in the sample.[6] We study the three most-traded currency pairs on the EBS system: euro-dollar, dollar-yen, and euro-yen.

The quote data, at the one-second frequency, consist of the highest bid quote and the lowest ask quote on the EBS system in these currency pairs, from which we construct one-second mid-quote series and compute one-minute exchange rate returns; all the quotes are executable and therefore represent the true price at that moment. The transactions data are at the one-minute frequency and provide detailed information on the volume and direction of trades that can be attributed to computers and humans in each currency pair. Specifically, the transactions volume data are broken down into categories specifying the "maker" and "taker" of the trades (i.e., human or computer), and the direction of the trades (i.e., buy or sell the base currency), for a total of eight different combinations. That is, the first transaction category may specify, say, the minute-by-minute volume of trade that results from a human taker buying the base currency by "hitting" a quote posted by a human maker. We would record this activity as the human-human buy volume, with the aggressor (taker) of the trade buying the base currency. The human-human sell volume is defined analogously, as are the other six buy and sell volumes that arise from the remaining combinations of computers and humans acting as makers and takers.

From these eight types of buy and sell volumes, we can construct, for each minute, trading volume and order flow measures for each of the four possible pairs of human and computer makers and takers: human-maker/human-taker (HH), computer-maker/human-taker (CH), human-maker/computer-taker (HC), and computer-maker/computer-taker (CC).[7] That is, the *sum* of the buy and sell volumes for each pair gives the volume of trade attributable to that particular combination of maker and taker (which we symbolize as, $Vol(HH)$ or $Vol(HC)$, for example). The *difference* between the buy and sell volume for each pair gives us the order flow attributable to that maker-taker combination (which we symbolize simply as HH or HC, for example). The sum of the four volumes, $Vol(HH + CH + HC + CC)$, gives the total volume of trade in the market. The sum of the four order flows, $HH + CH + HC + CC$, gives the total (market-wide) order flow.[8] Throughout the paper, we will use the expression "order flow" to refer both to the market-wide order flow and to the order flows from other possible decompositions, with the distinction clearly indicated. Importantly, the data allow us to consider order flow broken down by the type of trader who initiated the

[6] We do not use December 2007 in the sub-sample to avoid the influence of year-end effects.

[7] The naming convention for "maker" and "taker" reflects the fact that the "maker" posts quotes before the "taker" chooses to trade at that price. Posting quotes is, of course, the traditional role of the market-"maker."

[8] There is a very high correlation in this market between trading volume per unit of time and the number of transactions per unit of time, and the ratio between the two does not vary much over our sample. Order flow measures based on amounts transacted and those based on number of trades are therefore very similar.

trade, human-taker order flow ($HH + CH$) and computer-taker order flow ($HC + CC$).

The main goal of this paper is to analyze the effect algorithmic trading has on price discovery and volatility in the foreign exchange market. In our exchange rate data as in other financial data, the net of signed trades from the point of view of the takers (the market-wide order flow) is highly positively correlated with exchange rate returns, so that the takers are considered to be more "informed" than the makers. Thus, in our analysis of the relative effects of human and computer trades in the market, we consider prominently the order flow decomposition into human-taker order flow and computer-taker order flow. However, we also consider two other decompositions in our work. We consider the most disaggregated decomposition of order flow (HH, CH, HC, CC), as this decomposition allows us to study whether the liquidity suppliers, who are traditionally assumed to be "uninformed", are posting quotes strategically. This situation is more likely to arise in our data, which comes from a pure limit order book market, than in data from a hybrid market like the NYSE, because, as Parlour and Seppi (2008) point out, the distinction between liquidity supply and liquidity demand in limit order books is blurry.[9] We also decompose the data by maker type (human or computer) in order to study whether computers or humans are providing liquidity during the release of public information, which are periods of high exchange rate volatility and, often, market stress.

In our analysis, we exclude data collected from Friday 17:00 through Sunday 17:00 New York time from our sample, as activity on the system during these "non-standard" hours is minimal and not encouraged by the foreign exchange community. We also drop certain holidays and days of unusually light volume: December 24-December 26, December 31-January 2, Good Friday, Easter Monday, Memorial Day, Labor Day, Thanksgiving and the following day, and July 4 (or, if this is on a weekend, the day on which the U.S. Independence Day holiday is observed).

We show summary statistics for the one-minute returns and order flow data in Table 1. This table contains a number of noteworthy features. First, order flow, whether in total, broken down by human and computer takers, or broken down into the 4 possible pairs of makers and takers, is serially positively correlated, which is consistent with some informed trading models. For example, Easley and O'Hara (1987) model a situation where sequences of large purchases (sales) arise when insiders with positive (negative) signals are present in the market. He and Wang (1995) also show that insiders with good (bad) news tend to buy (sell) repeatedly until their private information is revealed in the prices. The positive serial correlation in order flow is also consistent with strategic order splitting, i.e. a trader willing to buy for informational or non-informational reasons and splitting his order to reduce market impact. Second, the standard deviations of the various order flows differ by exchange rates, by type of taker and across maker/taker pairs. These differences will be

[9] Parlour and Seppi (2008) note that in a limit order book investors with active trading motives, some of which are "informed" traders, may choose to post limit orders that are more aggresive than those a disinterested liquidity provider would use but less aggresive than market orders.

important in the interpretation of the upcoming VAR analysis and variance decompositions.

We show in Figure 1, from 2003 through 2007 for our three major currency pairs, the fraction of trading volume where at least one of the two counterparties was an algorithmic trader, i.e. $Vol(CH + HC + CC)$ as a fraction of total volume.[10] From its beginning in 2003, the fraction of trading volume involving AT grew by the end of 2007 to near 60% for euro-dollar, and dollar-yen trading, and to about 80% for euro-yen. Figure 2 shows, for our three currency pairs, the evolution over time of the four different possible types of trades (i.e. $Vol(HH)$, $Vol(CH)$, $Vol(HC)$, and $Vol(CC)$, as fractions of the total volume). By the end of 2007, in the euro-dollar and dollar-yen markets, human to human trades, in black, accounted for slightly less than half of the volume, and computer to computer trades, in green, for about ten to fifteen percent. In euro-dollar and dollar-yen, we note that $Vol(HC)$ and $Vol(CH)$ are about equal to each other, i.e. computers "take" prices posted by humans, in red, about as often as humans take prices posted by market-making computers, in blue. The story is different for the cross-rate, the euro-yen currency pair. By the end of 2007, there were more computer to computer trades than human to human trades. But the most common type of trade was computers trading on prices posted by humans. We believe this reflects computers taking advantage of short-lived triangular arbitrage opportunities, where prices set in the euro-dollar and dollar-yen markets are very briefly out of line with the euro-yen cross rate. In interpreting our results later in the paper, we will keep in mind that trading volume is largest in the euro-dollar and dollar-yen markets, and that price discovery happens mostly in those markets, not in the cross-rate. Our conclusions based on the euro-dollar and dollar-yen markets will then be more easily generalized than those based on the euro-yen market. Table 2 tabulates the averages of the volume fractions shown in Figures 1 and 2, both for the full 2006-2007 sample and the shorter three-month sub-sample.

3 How Correlated Are Algorithmic Trades and Strategies?

We first investigate the proposition that computers tend to have trading strategies that are more correlated than those of humans. Since the outset of the financial turmoil in the summer of 2007, articles in the financial press have suggested that AT programs tend to be similarly designed, leading them to take the same side of the market in times of high volatility and potentially exaggerating market movements.[11]

One such instance may have happened on August 16, 2007, a day of very high volatility in the dollar-yen market. On that day, the Japanese yen appreciated sharply against the U.S. dollar around 6:00 a.m. and 12:00 p.m. (NY time), as shown in Figure 3. The figure also shows, for each 30-minute interval in the day, computer-taker order flow ($HC + CC$) in the top panel and human-taker order flow ($HH + CH$) in the lower panel.

[10] The data in Figures 1 and 2 are 50-day moving averages of daily values, highlighting the broad trends over time.

[11] See, for instance, "Algorithmic Trades Produce Snowball Effects on Volatility," Financial Times, December 5, 2008.

The two sharp exchange rate movements mentioned happened when computers, as a group, aggressively sold dollars and bought yen. We note that computers, during these episodes, mainly traded with humans, not with other computers. Human order flow at those times was, in contrast, quite small, even though the overall trading volume initiated by humans (not shown) was well above that initiated by computers (human takers were therefore selling and buying dollars in almost equal amounts). The "taking" orders generated by computers during those time intervals were far more correlated than the taking orders generated by humans. After 12:00 p.m., human traders, as a whole, then began to buy dollars fairly aggressively, and the appreciation of the yen against the dollar was partially reversed. This is only a single example, of course, but it leads us to ask how correlated computer trades and strategies have tended to be overall.

We do not know precisely the exact mix of the various strategies used by algorithmic traders on EBS. Traders keep the information about their own strategies confidential, including, to some extent, from EBS, and EBS also keeps what they know confidential.[12] However, one can get a general sense of the market and of the strategies in conversations with market participants. About half of the algorithmic trading volume on EBS is believed to come from what is often known as the "professional trading community," which primarily refers to hedge funds and commodity trading advisors (CTAs). These participants, until very recently, could not trade manually on EBS, so all their trades were algorithmic. Some hedge funds and CTAs seek to exploit short-lived arbitrage opportunities, including triangular arbitrage, often accessing several trading platforms. Others implement lower-frequency strategies, often grouped under the statistical arbitrage appellation, including carry trades, momentum trades, and strategies spanning several asset classes. Only a very small fraction of the trading volume in our sample period is believed to have been generated by algorithms designed to quickly react to data releases. The other half (approximately) of the algorithmic trading volume comes from foreign exchange dealing banks, the only participants allowed on the EBS system until 2003. Some of the banks' algorithmic trading is clearly related to activity on their own customer-to-dealer platforms, to automate hedging activity, and to minimizing the impact of the execution of large orders. But a sizable fraction is believed to be proprietary trading implemented algorithmically, likely using a mix of strategies similar to those employed by hedge funds and CTAs. Overall, market participants generally believe that the mix of algorithmic strategies used in the foreign exchange market differs from that seen in the equity market, where optimal execution algorithms are thought to be relatively more prevalent.

The August 16, 2007 episode shown above was widely viewed as the result of a sudden unwinding of the yen-carry trade, with hedge funds and proprietary trading desks at banks rushing to close risky positions and buying yen to pay back low-interest loans. The evidence in this case raises the possibility that many

[12] EBS requires that new algorithmic traders on its system first test their algorithms in simulated conditions. EBS then routinely monitors the trading practices of its customers. A high number of excessively short-lived quotes (flashing) is discouraged, as is a very low ratio of trades to quotes.

algorithmic traders were using fairly similar carry trade and momentum strategies at the time, leading to the high correlation of algorithmic orders and to sharp exchange rate movements. Of course, this is only one episode in our two-year sample. Furthermore, episodes of very sharp appreciation of the yen due to the rapid unwinding of yen carry trades have occurred on several occasions since the late 1990s, some obviously before algorithmic trading was allowed in the market. The sharp move of the yen in October 1998, including a 1-day appreciation of the yen against the dollar of about 7 percent, is the best-known example of such an episode. Next, we investigate whether there is evidence that, over the entire sample, the strategies used by algorithmic traders have tended to be more correlated than those used by human traders.

If computers and humans are indifferent between taking or making liquidity at a given point in time, then we should observe that computers and humans trade with each other in proportion to their relative presence in the market. If, on the other hand, computers tend to have more homogeneous trading strategies, we should observe computers trading less among themselves and more with humans. At the extreme, if all computers used the very same algorithms and had the exact same speed of execution, we would observe no trading volume among computers. Therefore, the fraction of trades conducted between computers contains information on how correlated their strategies are.[13]

To investigate the proposition that computers tend to have trading strategies that are more correlated than those of humans we pursue the following approach. We first consider a simple benchmark model that assumes random and independent matching of traders. This model allows us to determine the theoretical probabilities of the four possible trades: Human-maker/human-taker, computer-maker/human-taker, human-maker/computer-taker and computer-maker/computer-taker. We then make inferences regarding the diversity of computer trading strategies based on how the trading pairs we observe compare to those the benchmark model predicts.

In the benchmark model there are H_m potential human-makers (the number of humans that are standing ready to provide liquidity), H_t potential human-takers, C_m potential computer-makers, and C_t potential computer-takers. For a given period of time, the probability of a computer providing liquidity to a trader is equal to $Prob(computer - make) = \frac{C_m}{C_m+H_m}$, which we label for simplicity as α_m, and the probability of a computer taking liquidity from the market is $Prob(computer - take) = \frac{C_t}{C_t+H_t} = \alpha_t$. The remaining makers and takers are humans, in proportions $(1 - \alpha_m)$ and $(1 - \alpha_t)$, respectively. Assuming that these events are independent, the probabilities of the four possible trades, human-maker/human-taker, computer-

[13] Stoffman (2007) uses a similar method to estimate how correlated individual investor strategies are compared to institutional investor strategies.

maker/human-taker, human-maker/computer-taker and computer-maker/computer taker, are:

$$\begin{aligned} Prob(HH) &= (1-\alpha_m)(1-\alpha_t) \\ Prob(HC) &= (1-\alpha_m)\alpha_t \\ Prob(CH) &= \alpha_m(1-\alpha_t) \\ Prob(CC) &= \alpha_m\alpha_t. \end{aligned}$$

These probabilities yield the following identity,

$$Prob(HH) \times Prob(CC) \equiv Prob(HC) \times Prob(CH),$$

which can be re-written as,

$$\frac{Prob(HH)}{Prob(CH)} \equiv \frac{Prob(HC)}{Prob(CC)}.$$

We label the first ratio, $RH \equiv \frac{Prob(HH)}{Prob(CH)}$, the "human-taker" ratio and the second ratio, $RC \equiv \frac{Prob(HC)}{Prob(CC)}$, the "computer-taker" ratio. In a world with more human traders (both makers and takers) than computer traders, each of these ratios will be greater than one, because $Prob(HH) > Prob(CH)$ and $Prob(HC) > Prob(CC)$ i.e., computers take liquidity more from humans than from other computers, and humans take liquidity more from humans than from computers. However, under the baseline assumptions of our random-matching model, the identity shown above states that the ratio of ratios, $R \equiv \frac{RC}{RH}$, will be equal to one. In other words, humans will take liquidity from other humans in a similar proportion that computers take liquidity from humans.

Turning to the data, under the assumption that potential human-takers are randomly matched with potential human-makers, i.e., that the probability of a human-maker/human-taker trade is equal to the one predicted by our model, $Prob(HH) = \frac{H_m \times H_t}{(H_m+C_m)\times(H_t+C_t)}$, we can now derive implications from observations of R, our ratio of ratios. In particular, finding $R > 1$ must imply that algorithmic strategies are more correlated than what our random matching model implies. In other words, for $R > 1$ we must observe that either computers trade with each other less than expected ($Prob(CC) < \frac{C_m \times C_t}{(H_m+C_m)\times(H_t+C_t)}$) or that computers trade with humans more than expected (either $Prob(CH) > \frac{C_m \times H_t}{(H_m+C_m)\times(H_t+C_t)}$ or $Prob(HC) > \frac{H_m \times C_t}{(H_m+C_m)\times(H_t+C_t)}$).

Our dataset allows us to estimate an ex-post proxy for R. Namely, for each trading day we estimate $\widehat{RH} = \frac{Vol(HH)}{Vol(CH)}$ and $\widehat{RC} = \frac{Vol(HC)}{Vol(CC)}$, where $Vol\,(HH)$ is the daily trading volume between human makers and human takers, and so forth. In Table 3 we show the mean of the daily ratio of ratios, $\widehat{R} = \frac{\widehat{RC}}{\widehat{RH}}$, for each currency pair for the full sample and the three-month sub-sample. In contrast to the above theoretical prediction that $R \equiv \frac{RC}{RH} = 1$, we find that for all currency pairs $\widehat{R}$ is statistically greater than one. This result is very robust: in euro-dollar, all daily observations of $\widehat{R}$ are above one, and only a very small fraction of the daily observations are below one for the other currency pairs. The results thus show that computers do not trade with each other as much as random matching would predict. We take this as evidence that algorithmic strategies are likely less diverse than the trading strategies used by human traders.

This finding, combined with the observed growth in algorithmic trading over time, may raise some concerns about the impact of AT on volatility in the foreign exchange market. As mentioned previously, some analysts have pointed to the possible danger of having many algorithmic traders take the same side of the market at the same moment. However, it is not a foregone conclusion that a high correlation of algorithmic strategies should necessarily lead to higher volatility or large swings in exchange rates. Both the high correlation of trading strategies and the widespread use of de-stabilizing strategies may need to be present to cause higher volatility. For instance, if many algorithmic traders use similar triangular arbitrage strategies, the high correlation of those strategies should have little impact on volatility, and may even lower volatility as it improves the efficiency of the price discovery process. Strategies designed to minimize the price impact of trades should also, a priori, not be expected to increase volatility. In contrast, if the high correlation reflects a large number of algorithmic traders using the same carry trade or momentum strategies, as in the August 2007 example shown at the beginning of this section, then there may be some reasons for concern. However, as noted earlier, episodes of sharp movements in exchange rates similar to that example have occurred in the past on several occasions, including well before the introduction of algorithmic trading in the foreign exchange market, suggesting that such episodes are a result of the dramatic unwinding of certain trading strategies, regardless of whether these strategies are implemented through algorithmic trading or not. In the next section, we explicitly investigate the relationship between the presence of algorithmic trading and market volatility.

4 The impact of algorithmic trading on volatility

In this section, we study whether the presence of algorithmic trading is associated with disruptive market behavior in the form of increased volatility. In particular, taking into account the potential endogeneity of algorithmic trading activity, we test for a causal relationship between the fraction of daily algorithmic trading relative to the overall daily volume, and daily realized volatility.

4.1 A first look

We first take an informal look at the data. Figure 4 shows monthly observations of annualized realized volatility (based on 1-minute returns) and of the fraction of algorithmic trading (the fraction of total trading volume involving at least one computer trader) for each of our currency pairs. As discussed earlier, there is a clear upward trend in the fraction of AT in the three currency pairs over 2006 and 2007. Realized volatility in euro-dollar, dollar-yen, and euro-yen declines slightly until mid-2007, and then rises in the second half of 2007, particularly sharply in the yen exchange rates, as the financial crisis begins.

In Figure 5, we study whether days with high market volatility are also days with a higher-than-usual fraction of algorithmic trading, and vice-versa. Using daily observations, we first sort the data into increasing deciles of realized volatility (the decile means are shown as bars in the graphs on the left).[14] We then calculate the mean fraction of AT for the days in each of these deciles (shown as lines in the same graphs). To account for the sharp upward trend in algorithmic participation over our sample, the daily fraction of algorithmic trading is normalized: we divide it by a 20-day moving average centered on the chosen observation (a moving average from day $t - 10$ through day $t + 10$, excluding day t). Next, we repeat the exercise, now sorting the daily data into increasing deciles of the normalized fraction of AT (the decile means are shown as bars in the graphs on the right) and calculating mean realized volatility for the days in each of these deciles (shown as lines in the same graphs). The results in Figure 5 (both the graphs on the left and the graphs on the right) show little or no relationship between the level of realized volatility on a particular day and the normalized fraction of AT on that same day. The highest decile in the euro-dollar currency pair may be the only possible exception, with a slight uptick evident in both volatility and AT activity. Finally, we note that, in untabulated results, for each of the three currency pairs, not one of the top 10 days in realized volatility is associated with a top ten day in the share of (normalized) AT.

The simple analysis in Figure 5 does not point to any substantial systematic link between AT activity and volatility. However, this analysis ignores the possible, and likely, endogeneity of algorithmic activity with regards to volatility, and therefore does not address the question of whether there is a causal relationship between algorithmic trading and volatility. In the remainder of this section, we attempt to answer this question through an instrumental variable analysis.

4.2 Identification

The main challenge in identifying a causal relationship between algorithmic trading and volatility is the potential endogeneity of algorithmic trading. That is, although one may conjecture that algorithmic trading

[14] With 498 daily observations, the first 9 deciles each include 50 observations, and the highest decile contains 48 observations.

impacts volatility, it is also plausible that algorithmic trading activity may be a function of the level of volatility. For instance, highly volatile markets may present comparative advantages to automated trading algorithms relative to human traders, which might increase the fraction of algorithmic trading during volatile periods. In contrast, however, one could also argue that a high level of volatility might reduce the informativeness of historical price patterns on which some trading algorithms are likely to base their decisions, and thus reduce the effectiveness of the algorithms and lead them to trade less. Thus, one can not easily determine in what direction the bias will go in an OLS regression of volatility on the fraction of algorithmic trading. To deal with the endogeneity issue, we adopt an instrumental variable (IV) approach as outlined below.

We are interested in estimating the following regression equation,

$$RV_{it} = \alpha_i + \beta_i AT_{it} + \gamma_i' \tau_{it} + \sum_{k=1}^{22} \delta_i RV_{it-k} + \epsilon_{it}, \tag{1}$$

where $i = 1, 2, 3$ represents currency pairs and $t = 1, ..., T$, represents time. RV_{it} is (log) realized daily volatility, AT_{it} is the fraction of algorithmic trading at time t in currency pair i, τ_{it} is either a time trend or a set of time dummies that control for secular trends in the data, and ϵ_{it} is an error term that is assumed to be uncorrelated with RV_{it-k}, $k \geq 1$, but not necessarily with AT_{it}. The large number of lags of volatility, which covers the business days of the past month, is included to control for the strong serial correlation in volatility (e.g. Andersen, Bollerslev, Diebold, and Labys, 2003 and Bollerslev and Wright, 2000). The exact definitions of RV_{it}, AT_{it}, and τ_{it} are given below.

The main focus of interest is the parameter β_i, which measures the impact of algorithmic trading on volatility in currency pair i. However, since AT_{it} and ϵ_{it} may be correlated, due to the potential endogeneity discussed above, the OLS estimator of β_i may be biased. In order to obtain an unbiased estimate, we will therefore consider an instrumental variable approach. Formally, we need to find a variable, or set of variables, z_{it}, that is uncorrelated with ϵ_{it} (validity of the instrument) and correlated with AT_{it} (relevance of the instrument).

The instrument we propose to use is the fraction of trading floors equipped to trade algorithmically on EBS relative to the total number of trading floors linked to the EBS system.[15] That is, in order to place algorithmic trades on EBS, a special user interface is required, and the total number of trading floors with such user interfaces thus provides a measure of the overall algorithmic trading "capacity" in the market. The ratio of these algorithmic trading floors to the total number of trading floors provides a measure of the

[15] More precisely, we actually observe a time series of the number of EBS "deal codes" of each type over our sample period. Generally speaking, EBS assigns a deal code to each trading floor equipped with at least one of its terminals, and records whether they are equipped to trade algorithmically or not. These data are confidential.

potential fraction of algorithmic trading. Since setting up an algorithmic trading operation likely takes several months, the number of trading floors of each type is clearly exogenous with regards to daily market volatility; the fraction of AT trading floors is therefore a valid instrument. In addition, it is positively correlated with the fraction of algorithmic trading, and it provides a relevant instrument as seen from the tests for weak instruments discussed below.

Under the breakdown provided by EBS, there are three types of trading floors linked to the EBS system: purely algorithmic trading floors, purely manual trading floors, and dual trading floors, those equipped to handle both manual and algorithmic trades. We consider two natural instrumental variables: the fraction of pure AT trading floors over the total number of trading floors (including pure AT, manual, and dual ones), and the fraction of the sum of pure AT *and* dual trading floors over the total number. Since it is not obvious which variable is the better instrument, we use both simultaneously.[16]

The data on AT trading floors are provided on a monthly basis, whereas the data on realized volatility and algorithmic trading are sampled on a daily frequency. We therefore transform the trading floor data to daily data by repeating the monthly value each day of the month. Although this leads to a dataset of two years of daily data, the number of daily observations (498) overstates the effective number of observations, since the coefficient on AT participation will be identified from monthly variations in the instrumental variables. Transforming the instruments to a daily frequency is, however, more efficient than transforming all data to a monthly frequency, since the daily data help to identify the monthly shifts.

The instrumental variable regressions are estimated using Limited Information Maximum Likelihood (LIML), and we test for weak instruments by comparing the first stage $F-$statistic for the excluded instruments to the critical values of Stock and Yogo's (2005) test of weak instruments. We use LIML rather than two-stage least squares since Stock and Yogo (2005) show that the former is much less sensitive to weak instruments than the latter (see also Stock et al., 2002).

4.3 Variable definitions

4.3.1 Realized Volatility

Volatility is measured as the *daily realized volatility* obtained from one minute returns; that is, the volatility measure is equal to the square root of the daily sum of squared one minute log-price changes. The use of realized volatility, based on high-frequency intra-daily returns, as an estimate of ex-post volatility is now well established and generally considered the most precise and robust way of measuring volatility. Although

[16] Regressions not reported here show that using the fraction of pure AT trading floors as a single instrument gives qualitatively similar results to those presented below based on both instruments. Using the fraction of the sum of both pure and dual AT trading floors as a single instrument also leads to the same qualititative conclusion, but with more signs of weak instruments.

many older studies relied on five minute returns in order to avoid contamination by market microstructure noise (e.g. Andersen et al., 2001), recent work shows that sampling at the one-minute frequency, or even higher frequencies, does not lead to biases in liquid markets (see, for instance, the results for liquid stocks in Bandi and Russel, 2006, and the study by Chaboud et al., 2007, who explicitly examine EBS data on the euro-dollar exchange rate during 2005 and finds that sampling frequencies upwards of once every 20 seconds does not lead to noticeable biases). Here, we restrict ourselves to using minute-by-minute data.[17] Following the common conventions in the literature on volatility modelling (e.g. Andersen, Bollerslev, Diebold, and Labys, 2003), the realized volatility is log-transformed to obtain a more well behaved time-series.

4.3.2 Algorithmic trading

We consider two measures of the fraction of algorithmic trading, AT_{it}, in a given currency pair: the computer-participation fraction and the computer-taker fraction. The first is simply the percent of the overall trading volume that includes an algorithmic trader as either a maker or a taker ($Vol(CH + HC + CC)$); that is, the percent of trading volume where a computer is involved in at least one side of the trade. In addition, we also consider an alternative measure defined as the fraction of overall trading volume that is due to a computer-taker ($Vol(HC + CC)$).

4.3.3 Time controls

As seen in Figure 4, there is a clear secular trend in the computer-participation fraction,[18] which is not present in realized volatility. Euro-dollar, dollar-yen, and euro-yen volatility is trending down at the beginning of the period and starts to trend up in the summer of 2007. In order to control for the trend in algorithmic trading in the regression, we include either a "linear quarterly" time trend or a full set of year-quarter dummies, one for each year-quarter pair in the data (8 dummies). That is, the linear quarterly time trend stays constant within each quarter and increases by the same amount each quarter, whereas the year-quarter dummies allows for a more flexible trend specification that can shift in arbitrary fashion from year-quarter to year-quarter. Both secular trend specifications are thus fixed within each quarter. This restriction is imposed in order to preserve the identification coming from the monthly instrumental variables. Using monthly, or finer, time dummies would eliminate the variation in the instrument and render the model unidentified. Although it is theoretically possible to include a monthly time trend, this would lead to very weak identification empirically.

[17] Using realized volatility based on five-minute returns leads to results that are very similar to those reported below for the one-minute returns, and the qualitative conclusions are identical.

[18] The same is true for the computer-taker fraction, not shown in the figure.

4.4 Empirical results

The regression results are presented in Table 4. We present OLS and LIML-IV results, with either the quarterly trend or the year-quarter dummies included. We show in Panels A and B the results for the computer-participation volume, and in Panels C and D the results for computer-taker volume. We report results for the sample starting in January 2006 and ending in December 2007. In order to save space, we only show the estimates of the coefficients in front of the fraction of algorithmic trading volume variables.

The OLS results, which are likely to be biased due to the aforementioned endogeneity issues, show a fairly clear pattern of a positive correlation between volatility and AT participation, with several positive and statistically significant coefficients. The R^2s are fairly large, reflecting the strong serial correlation in realized volatility, which is picked up by the lagged regressors. There are also no systematic differences between the quarterly trend and quarterly dummies specifications.

Turning to the more interesting IV results, which control for the endogeneity bias, the coefficient estimates change fairly dramatically. All point estimates are now negative and some of them are statistically significant. Thus, if there is a *causal* relationship between the fraction of algorithmic trading and the level of volatility, all evidence suggests that it is negative, such that increased AT participation lowers the volatility in the market. The stark difference between the IV and OLS results shows the importance of controlling for endogeneity when estimating the causal effect of AT on volatility; the opposite conclusion would have been reached if one ignored the endogeneity issue. The evidence of a statistically significant relationship is fairly weak, however, with most coefficients statistically indistinguishable from zero. The more restrictive quarterly trend specification suggests a significant relationship for the euro-dollar and dollar-yen, but this no longer holds if one allows for year-quarter dummies.

To the extent that the estimated coefficients are statistically significant, it is important to discuss the *economic* magnitude of the estimated relationship between AT and volatility. The regression is run with log volatility rather than actual volatility, which makes it a little less straightforward to interpret the size of the coefficients. However, some back-of-the-envelope calculations can provide a rough idea. Suppose that the coefficient on computer participation is about -0.01, which is in line with the coefficient estimates for the euro-dollar. The average monthly shift in computer participation in the euro-dollar is about 1.5 percentage points and the average log-volatility in the euro-dollar is about 3.76 (with returns calculated in basis points), which implies an annualized volatility of about 6.82 percent. Increasing the computer participation fraction by 1.5 percentage points decreases log-volatility by 0.015 and results in an annualized volatility of about 6.72. Thus, a typical change in computer participation might change volatility by about a tenth of a percentage point in annualized terms, a small effect.

The first stage $F-$statistics for the excluded instruments in the IV regressions are also reported in Panels B and D. Stock and Yogo (2005) show that this $F-$statistic can be used to test for weak instruments. Rejection of the null of weak instruments indicates that standard inference on the IV-estimated coefficients can be performed, whereas a failure to reject indicates possible size distortions in the tests of the LIML coefficients. The critical values of Stock and Yogo (2005) are designed such that they indicate a maximal actual size for a nominal sized five percent test on the coefficient. Thus, in the case considered here with two excluded instruments and one endogenous regressor, a value greater than 8.68 for this $F-$statistic indicates that the maximal size of a nominal 5 percent test will be no greater than 10 percent, which might be deemed acceptable; a value greater than 5.33 for the $F-$statistic indicates a maximal size of 15 percent for a nominal 5 percent test. In general, the larger the $F-$statistic, the stronger the instruments. As is evident from the table, there are no signs of weak instruments in the specification with a quarterly trend. There are, however, signs of weak instruments in the case with year-quarter dummies, for the euro-yen. This is not too surprising given that the instruments only change on a monthly frequency, and the year-quarter dummies therefore put a great deal of strain on the identification mechanism. Importantly, though, the results for the two major currency pairs are robust to any weak-instrument problems and the reported coefficients and standard errors are unbiased.

To sum up, the evidence of any causal effect of algorithmic trading on volatility is not strong, but what evidence there is points fairly consistently towards a negative relationship. There is thus no systematic statistical evidence to back the often-voiced opinion that AT leads to increased levels of market volatility. If anything, the contrary appears to be true.

5 Who provides liquidity during the release of public announcements?

In the previous section we discuss one of the major concerns regarding algorithmic trading, namely, whether AT causes exchange rate volatility. We now examine another major concern, whether AT improves or reduces liquidity during stress periods, when it is arguably needed the most. To answer this question, we cannot simply regress computer-maker volume, a proxy for liquidity provided by computers, on exchange rate volatility, a proxy for stress periods, because, as we discussed in the previous section, algorithmic volume and volatility are endogenous variables. In contrast to the previous section we do not estimate an IV regression, as there are no obvious instruments for volatility.[19] Instead, we follow the event study literature

[19] One could consider macroeconomic news announcements as potential instruments for volatility. However, macroeconomic news announcements are exogeneous variables that cause both foreign exchange rate volatility and liquidity changes. Since we cannot assume that the effect macroeconomic news announcements have on liquidity is only due to the effect macroeconomic

and compare the liquidity provision by humans and computers during U.S. nonfarm payroll announcements, a period of exogenously heightened volatility, to the liquidity provision by both types of agents during non-announcement days. This comparison will help us determine who provides relatively more liquidity during stress periods. We note that, when we consider liquidity provision by humans and computers following other important macroeconomic news announcements, the results are qualitatively similar. However, we focus in this section on the nonfarm payroll announcement only, as it routinely generates the highest volatility of all US macroeconomic announcements.[20]

We consider two liquidity provision estimates: a one-minute estimate and a one-hour estimate. The one-minute estimate is calculated using volume observations from 8:30 a.m. to 8.31 a.m. ET (when U.S. nonfarm payroll is released), while the one-hour estimate is calculated using observations from 8:25 am to 9:24 am ET. We define the one-minute (one-hour) liquidity provision by humans, LH, as the sum of human-maker volume, $Vol(HH+HC)$, divided by total volume during that period, and the one-minute (one-hour) liquidity provision by computers, LC, as the sum of computer-maker volume, $Vol(CC+CH)$, divided by total volume during that period. Similar to the liquidity provision measures, we define the one-minute volatility as the squared 1-minute return from 8:30 a.m. to 8.31 a.m. ET and the one-hour volatility as the sum of squared 1-minute returns from 8:25 am to 9:24 am ET.

To compare liquidity provision by humans and computers during announcement times to liquidity provision during (more tranquil) non-announcement times, we could estimate the average liquidity provision during announcement times and compare it to the average liquidity provision during non-announcement times, with both means taken over the entire sample period. However, as we discussed previously, exchange rate trading volumes and the shares of liquidity provision by humans and computers exhibit clear trends over our sample, making the comparison of the two different means problematic. Alternatively, and this is the methodology we follow, on each announcement day we estimate the ratio of liquidity provision on that day relative to the liquidity provision on days surrounding the announcement. This amounts to using a non-parametric approach to detrend the data. The time series of these ratios will be stationary, and we can then test the hypothesis that the ratio is greater than one. Specifically, we divide the one-minute (one-hour) liquidity provision by humans, LH_a, and computers, LC_a, estimated on announcement day t by the one-minute (one-hour) liquidity provision by humans, LH_n, and computers, LC_n, respectively, estimated during the surrounding non-announcement day period, defined as 10 business days before and after a nonfarm payroll release date t. The liquidity provision measures on the non-announcement days are calculated in the same manner as on the announcement days, using data only for the periods 8:30 a.m. to 8.31 a.m. ET or 8:25 am to 9:24

news announcements have on volatility, the exclusion restriction required by IV estimation is violated.

[20] Andersen and Bollerslev (1998), among others, refer to the nonfarm payroll report as the “king” of announcements, because of the significant sensitivity of most asset markets to its release.

am ET, for the one-minute and one-hour measures, respectively.[21] We follow the same procedure with our one-minute and one-hour volatility estimates.

Consistent with previous studies, we show in Table 5 Panel A that the one-hour volatility on nonfarm payroll announcement days is 3 to 6 times larger than during non-announcement days. The one-minute volatility is 15 to 30 times larger during announcement days compared to non-announcement days. As expected, given the fact that we focus on a U.S. data release, the volatility increase is smaller in the cross-rate, the euro-yen exchange rate, than in the euro-dollar and yen-dollar exchange rates. Focusing on the statistically significant estimates, we show in Table 5 Panel B that, as a share of total volume, human-maker volume tends to increase during the minute of the announcement (the one-minute ratio $\frac{LH_a}{LH_n}$ is greater than one), while computer-maker volume tends to decrease (the one-minute ratio $\frac{LC_a}{LC_n}$ is less than one). Interestingly, this pattern is reversed when we focus on the one-hour volume estimates for the euro-dollar and euro-yen exchange rate markets. In relative terms, computers do not increase their provision of liquidity as much as humans do during the minute following the announcement. However, computers increase their provision of liquidity relatively more than humans do over the entire hour following the announcement, a period when market volatility remains quite elevated.

We note that, over our sample period, the U.S. nonfarm payroll data releases were clearly the most anticipated and most influential U.S. macroeconomic data releases. They often generated a large initial sharp movement in exchange rates, followed by an extended period of volatility. The behavior of computer traders observed in the first minute could reflect the fact that many algorithms are not designed to react to the sharp, almost discrete, moves in exchange rates that often come at the precise moment of the data release. Some algorithmic traders may then prefer to pull back from the market a few seconds before 8:30 a.m. ET on days of nonfarm payroll announcements, resuming trading once the risk of a sharp initial price movement has passed. But the data show that algorithmic traders, as a whole, do not shrink back from providing liquidity during the extended period of volatility that follows the data releases.

6 Price Discovery

In the previous three sections, we analyze questions that are primarily motivated by practical concerns regarding algorithmic trading, such as whether computer traders induce volatility or reduce liquidity. In this section we turn to questions that are driven more by the market microstructure literature, but that also lead

[21] For simplicity, we label the 10 business days before and after the nonfarm payroll announcement as non-announcement days. However, during this 20-day period there are both days with no macroeconomic news and days with news. For instance, every Thursday, including the day before the monthly nonfarm payroll number is released, initial jobless claims are released. Thus, our estimation will likely be biased towards not finding statistically different behavior across the two periods. As we show in Table 5, volatility is, on average, much lower during this 20-day period than on nonfarm payroll days, and therefore the period still serves as a good benchmark.

to interesting practical insights regarding the effects and nature of algorithmic trading. In particular, we study price discovery within a vector autoregressive framework, which enables us to evaluate to what extent humans or computers represent the "informed" traders in the market. Our findings reveal several interesting features regarding the impact of algorithmic trades and the order placement behavior of computer traders.

6.1 Who are the "informed" traders, humans or computers?

We first investigate whether human or computer trades have a more "permanent" impact on prices. To this end, we estimate return-order flow dynamics in a structural vector autoregressive (VAR) framework in the tradition of Hasbrouck (1991a), where returns are contemporaneously affected by order flow, but order flow is not contemporaneously affected by returns. Similar to Hasbrouck's (1996) decomposition of program and nonprogram order flow, we decompose order flow into two components: human-taker $\left(OF^{(ht)} = HH + CH\right)$ and computer-taker $\left(OF^{(ct)} = HC + CC\right)$, and thus we estimate for each currency i one return equation and two order flow equations. In light of Evans and Lyons (2008) findings, we estimate the structural VAR with U.S. macroeconomic news surprises as exogenous variables that affect both returns and order flow. Specifically, we estimate the following system of equations for each currency i,

$$\begin{aligned}
r_{it} &= \alpha^r + \sum_{j=1}^{J} \beta^r_{ij} r_{it-j} + \sum_{j=0}^{J} \gamma^{rct}_{ij} OF^{(ct)}_{it-j} + \sum_{j=0}^{J} \gamma^{rht}_{ij} OF^{(ht)}_{it-j} + \sum_{k=1}^{K} \delta^r_{ik} S_{kt} + \varepsilon^r_{it}, \qquad (2)\\
OF^{(ht)}_{it} &= \alpha^{OF}_{ht} + \sum_{j=1}^{J} \beta^{OF}_{ijht} r_{it-j} + \sum_{j=1}^{J} \gamma^{OF^{(ht)}}_{ijht} OF^{(ht)}_{it-j} + \sum_{j=1}^{J} \gamma^{OF^{(ct)}}_{ijht} OF^{(ct)}_{it-j} + \sum_{k=1}^{K} \delta^{OF}_{ikht} S_{kt} + \varepsilon^{OF^{(ht)}}_{it},\\
OF^{(ct)}_{it} &= \alpha^{OF}_{ct} + \sum_{j=1}^{J} \beta^{OF}_{ijct} r_{it-j} + \sum_{j=1}^{J} \gamma^{OF^{(ct)}}_{ijct} OF^{(ct)}_{it-j} + \sum_{j=1}^{J} \gamma^{OF^{(ht)}}_{ijct} OF^{(ht)}_{it-j} + \sum_{k=1}^{K} \delta^{OF}_{ikct} S_{kt} + \varepsilon^{OF^{(ct)}}_{it}.
\end{aligned}$$

Here r_{it} is the 1-minute exchange rate return for currency i at time t; OF^{ht}_{it} is the currency i human-taker order flow at time t; OF^{ct}_{it} is the currency i computer-taker order flow at time t; and S_{kt} is the macroeconomic news announcement surprise for announcement k at time t defined as the difference between the announcement realization and its corresponding market expectation. We use Bloomberg's real-time data on the expectations and realizations of $K = 28$ U.S. macroeconomic fundamentals to calculate S_{kt}. The 28 announcements we consider are similar to those in Andersen et al. (2003, 2007) and Pasquariello and Vega (2007).[22] Since units of measurement vary across macroeconomic variables, we standardize the resulting surprises by dividing each

[22] Our list of U.S. macroeconomic news announcements is the same as the list of announcements in Andersen et al. (2007) and Pasquariello and Vega (2007) with the addition of three announcements: unemployment rate, core PPI and core CPI. Andersen et al. (2007) and Pasquariello and Vega (2007) use International Money Market Services (MMS) data on the expectations of U.S. macroeconomic fundamentals. In contrast, we use Bloomberg data because the MMS data are no longer available after 2003. Bloomberg provides survey data similar to those MMS previously provided.

of them by their sample standard deviation. Economic theory suggests that we should also include foreign macroeconomic news announcements in equation (2). However, previous studies find that exchange rates do not respond much to non-U.S. macroeconomic announcements, even at high frequencies (e.g. Andersen et al., 2003), so we expect the omitted variable bias in our specification to be small.

The underlying economic model is based on continuous time, and we thus estimate the VAR using the highest sample frequency available to us, minute-by-minute data. The estimation period is restricted to the $2006-2007$ sample, and the total number of observations for each currency pair is $717,120$ in the full sample and $89,280$ in the three-month sub-sample (September, October and November of 2007). In both samples, 20 lags are included in the estimated VARs, i.e. $J = 20$.

Our specification in equation (2) does not allow human-taker order flow to contemporaneously affect computer-taker order flow or vice-versa. The advantage of this approach is that we can estimate the impulse response functions without giving more importance to a particular type of order flow, i.e., we do not need to assume a particular ordering of the human-taker and computer-taker order flow in the VAR. The disadvantage is that the human-taker and computer-taker order flow shocks may not be orthogonal. However, in our estimation this does not appear to be a problem, as our residuals are found to be approximately orthogonal (the correlation between the human-taker and computer-taker equation residuals are -0.001, -0.1 and -0.1 for the euro-dollar, yen-dollar, and euro-yen exchange rates respectively). As a robustness check, we also estimate the VAR with two different orderings. We first assume human-taker order flow affects computer-taker order flow contemporaneously, and then assume the opposite ordering. This latter approach allows us to compute upper and lower bound impulse responses. These results are presented in the Appendix, and show that the results presented here are not sensitive to alternative identification schemes in the VAR.

Before considering the impulse response functions and the variance decompositions, we briefly summarize the main lessons from the estimated coefficients in the VAR. Focusing on the return equation, we find that minute-by-minute returns tend to be negatively serially correlated, with the coefficient on the first own lag varying between -0.08 and -0.15; there is thus some evidence of mean reversion in the exchange rates at these high frequencies, which is a well-know empirical finding. Both order flows are significant predictors of returns. The price impact of the lagged order flows range from around 4 to 18 basis points per billion units of order flow (denominated in the base currency), as compared to a range of approximately $28 - 100$ basis points in the contemporaneous order flow. As theory would predict, we find that U.S. macroeconomic news announcements affect less the euro-yen exchange rate (i.e., the R^2 of regressing the euro-yen exchange rate on macroeconomic news surprises and restricting the sample to announcement-only observations is 23%) than the euro-dollar and dollar-yen exchange rates (i.e., the R^2s of an announcement-only sample are 60% and 59%, respectively). However, U.S. macroeconomic news announcements still have an effect on the cross-rate to the extent that the U.S. economy is more or less correlated with the Japanese or the Euro-area economy.

Focusing on the order-flow equations, we find that the first own lag in both order flow equations is always highly significant, and typically around 0.1 for all currency pairs. There is thus a sizeable first-order autocorrelation in the human-taker and computer-taker order flows. The coefficients on the first order cross-lags in the order flow regressions are most often substantially smaller than the coefficient on the own lag and vary in signs. Lagged returns have a small but positive impact on order flow, suggestive of a form of trend chasing by both computers and humans in their order placement.

We note that despite the strongly significant estimates that are recorded in the VAR estimations, the amount of variation in the order flow and return variables that is captured by their lagged values is very limited. The R^2 for the estimated equations with only lagged variables are typically around three to ten percent for the order flow equations, and between one and three percent for the return equations. This can be compared to an R^2 of 20 to 30 percent when one includes contemporaneous order flow.

6.2 Impulse Response Function and Variance Decomposition Results

As originally suggested by Hasbrouck (1991b), we use the impulse response functions to assess the price impact of various order flow types, and the variance decompositions to measure the relative importance of the variables driving foreign exchange returns. In Table 6 Panel A, we show the results from the impulse response analysis based on the estimation of equation (2), using the full sample for 2006-2007 and the three-month sub-sample, when the size of the shock is the same across the different types of order flow: a one billion base currency shock to order flow. We also show the results when the size of the shock varies according to the average size shock: a one standard deviation base currency shock to order flow (Table 6 Panel B). We show both the short-run (instantaneous) impulse responses, the long-run cumulative responses, and the difference between the two responses. The long-run statistics are calculated after 30-minutes, at which point the cumulative impulse responses have converged and can thus be interpreted as the long-run total impact of the shock. All the responses are measured in basis points. The standard errors reported in the tables are calculated by bootstrapping, using 200 repetitions.

Starting with a hypothetical shock of one billion base currency order flow, the results in Table 6 Panel A, show that the immediate response of prices to human-taker order flow is often larger than the immediate response to computer-taker order flow. This may partially be attributed to the fact that some of the algorithmic trading is used for the optimal execution of large orders at a minimum cost. Algorithmic trades appear to be successful in that endeavor, with computers likely breaking up the larger orders and timing the smaller trades to minimize the impact on prices. We emphasize, though, that the differences in price impact, which range from 1 to 8 basis points, are not very large in economic terms. Furthermore, we find that the result can be reversed in the long-run and in the three-month sub-sample. For example, the euro-dollar human-taker price impact is larger than the computer-taker price impact in the short-run, but the opposite is true in the long-run and in the three-month sub-sample.

In contrast to these results, the response to a hypothetical one standard deviation shock to the different order flows (Table 6 Panel B) consistently shows that in the euro-dollar and dollar-yen markets, humans have a bigger impact on prices than computers and the differences are relatively large. For example, a one standard deviation shock to human-taker order flow in the yen-dollar exchange rate market has an average long-run effect of 0.9 basis points compared to an average effect of 0.3 basis points for computer-taker order flow. Interestingly, the difference in price impact in the cross-rate, the euro-yen exchange rate, is very small. In this market, computers have a clear advantage over humans in detecting and reacting more quickly to triangular arbitrage opportunities so that a large proportion of algorithmic trading contributes to more efficient price discovery. It is then not so surprising that in this market computers and humans, on average, appear to be equally "informed."

In Table 7 we report the fraction of the total (long-run) variance in returns that can be attributed to innovations in human-taker order flow and computer-taker order flow.[23] Following Hasbrouck (1991b), we interpret this variance decomposition as a summary measure of the informativeness of trades, and thus, in the current context, a comparison of the relative informativeness of the different types of order flow. Consistent with the results from the impulse response functions based on a one standard deviation shock to order flow, we find that in the euro-dollar and dollar-yen exchange rate markets human-taker order flow explains much more of the total variance in returns than computer-taker order flow. Specifically, human-taker order flow explains about 30 percent of the total variance in returns compared to only 4 percent explained by computer-taker order flow.

The fact that human-taker order flow explains a bigger portion of total variance in returns is not surprising because human-taker volume is about 75 percent of total volume in these two markets in the full sample period and about 65 percent of total volume in the three-month sub-sample (see Table 2). Moreover large buy (sell) orders tend to be human-taker orders, i.e. we show in Table 1 that the standard deviation of human-taker order flow is twice as big as that of the computer-taker order flow. But, do computers tend to contribute "disproportionately" little to the long-run variance in returns relative to their trading volume? To answer this question we do a back-of-the-envelope calculation. We compute the relative share of the *explained* variance that is due to computer-taker order flow as the percent of total variation in returns explained by computer-

[23] The variance decompositions are virtually identical in the short- and long-run and thus we only show the long-run decomposition results.

taker order flow divided by the percent of total variation in returns explained jointly by both human-taker and computer-taker order flow. For example, this relative share is $14\% = 100 \times \frac{4.74}{34}$ (Table 7) in the euro-dollar market. We can then compare this relative share to the fraction of overall trading volume that is due to computer-taker volume, which we show in Table 2. In the full 2006-2007 sample for the euro-dollar and the dollar-yen currency pairs, the fraction of volume due to computer-takers is about twice as large as the fraction of the explained long-run variance that is due to computer-taker order flow. In the euro-yen, the fractions are approximately equal. These results are fairly similar in the three-month sub-sample, although the fraction of explained variance has increased somewhat compared to the volume fraction. Thus, in the two major currency pairs, there is evidence that computer-taker order flow contributes relatively less to the variation in returns than one would infer from just looking at the proportion of computer-taker volume.

6.3 Are liquidity providers "uninformed"?

We now turn to examine whether liquidity providers post quotes strategically. To this end we augment equation (2) and decompose order flow into four components. Specifically, we estimate the following system of equations for each currency i,

$$
\begin{aligned}
r_{it} &= \alpha^r + \sum_{j=1}^{J} \beta_{ij}^r r_{it-j} + \sum_{l=1}^{L} \sum_{j=0}^{J} \gamma_{ij}^{rl} OF_{it-j}^{(l)} + \sum_{k=1}^{K} \delta_{ik}^r S_{kt} + \varepsilon_{it}^r, \qquad (3) \\
OF_{it}^{(l)} &= \alpha_l^{OF} + \sum_{j=1}^{J} \beta_{ijl}^{OF} r_{it-j} + \sum_{l=1}^{L} \sum_{j=1}^{J} \gamma_{ijl}^{OF^{(l)}} OF_{it-j}^{(l)} + \sum_{k=1}^{K} \delta_{ikl}^{OF} S_{kt} + \varepsilon_{it}^{OF^{(l)}}.
\end{aligned}
$$

where r_{it} is the 1-minute exchange rate return for currency i at time t; $L = 4$, $OF_{it}^{(1)} = OF_{it}^{HH}$ is the currency i human-maker/human-taker order flow at time t; $OF_{it}^{(2)} = OF_{it}^{CH}$ is the currency i computer-maker/human-taker order flow at time t; $OF_{it}^{(3)} = OF_{it}^{HC}$ is the currency i human-maker/computer-taker order flow at time t; $OF_{it}^{(4)} = OF_{it}^{CC}$ is the currency i computer-maker/computer-taker order flow at time t; S_{kt} is the macroeconomic news announcement surprise for announcement k at time t.[24]

In addition to identifying whether traders, on average, have a more permanent impact on prices when trading with humans than with computers, this specification also allows us to observe the effect order flow has on prices when, for instance, no party has a speed advantage, i.e. both parties are humans or both parties are computers, and when either the maker has a speed advantage, CH, or the taker has a speed advantage, HC. This distinction may be particularly useful when analyzing the cross-rate, where computers likely have

[24] In the Appendix, we analyze the robustness of this structural VAR by also estimating impulse responses and variance decompositions from all possible triangular identification schemes, only imposing that returns are ordered last in the VAR.

a clear advantage over humans in detecting short-lived triangular arbitrage opportunities.

Starting with a hypothetical shock of one billion base currency order flow, the results in Table 8 Panel A show that there is no clear pattern in which order flow impacts prices the most. However, the dynamics of the VAR system help reveal an interesting finding: There is a consistent and often large short-run over-reaction to CC and CH shocks. That is, as seen in Table 8, the short run response to a CC or CH order flow shock is always larger than the long-run response, and sometimes substantially so. The euro-dollar in the sample covering September, October, and November of 2007 provides an extreme case where the initial reaction to a one billion dollar CC shock is a 22 basis point move, but the long-run cumulative reaction is just 6 basis points. Interestingly, the opposite pattern is true for the HH order flow shocks, where there is always an initial *under*-reaction in returns. To the extent that an over-reaction of prices to order flow is suggestive of the presence of liquidity traders, these impulse response patterns suggest that computers provide liquidity when the probability of trading with an informed trader is low.[25]

The response to a hypothetical one standard deviation shock to the different order flows consistently shows that HH order flow has a bigger impact on prices than CC order flow (Table 8 Panel B) and that the differences are large. In particular, a one standard deviation shock to HH order flow has an average long-run effect of 0.6 basis points across currencies compared to a one standard deviation shock to CC order flow, which has an average effect of 0.1 basis points. Interestingly, we observe that when humans trade with other humans they influence prices more than when they trade with computers (the impact of HH on prices is bigger than the impact of CH on prices), and when computers trade with other computers they influence prices less than when they trade with humans (the impact of CC on prices is smaller than the impact of HC on prices). Our interpretation is that computers provide liquidity more strategically than humans, so that the counterparty cannot affect prices as much. This interpretation is consistent with the over-reaction of prices to CC and CH order flow described above: Computers appear to provide liquidity when adverse selection costs are low. This finding relates to the literature that proposes to depart from the prevalent assumption that liquidity providers in limit order books are passive.[26]

We also find that the price response to order flow varies across currencies as these markets differ along several dimensions. Trading volume is largest in the euro-dollar and dollar-yen markets, compared to the euro-yen market, and price discovery clearly happens mostly in the two largest markets. In the cross-rate

[25] Dynamic learning models with informed and uninformed investors predict that prices will temporarily over-react to uninformed order flow and under-react to informed order flow (e.g., Albuquerque and Miao, 2008). We note that the over- and under-reaction of prices to a particular type of order flow is different from the over- and under-reaction of prices to public news, which are both considered a sign of market inefficiency. Order flow types are not public knowledge, so that agents cannot trade on this information.

[26] For example, Chakravarty and Holden (1995), Kumar and Seppi (1994), Kaniel and Liu (2006), and Goettler, Parlour and Rajan (2007) allow informed investors to use both limit and market orders. Bloomfield, O'Hara and Saar (2005) argue that informed traders are natural liquidity providers and Angel (1994) and Harris (1998) show that informed investors can optimally use limit orders when private information is sufficiently persistent.

market, the euro-yen, computers have a speed advantage over humans in profiting from triangular arbitrage opportunities, where prices set in the euro-dollar and dollar-yen markets are very briefly out of line with the euro-yen rate. Consistent with this speed advantage we observe that human-maker/computer-taker order flow has a larger price impact in the cross-rate market than in the other two markets.

In addition to the impulse response functions, we also report the long-run forecast variance decomposition of returns in Table 9 for both the full sample and the three-month sub-sample. Consistent with the impulse response functions to a one standard deviation shock to order flow, the HH order flow makes up the dominant part of the variance share in the euro-dollar and dollar-yen exchange rate markets. In the last three months of the sample, this share has generally decreased. The share of variance in returns that can be attributed to the CC order flow is surprisingly small, especially in the latter sub-sample, given that this category of trades represents a sizeable fraction of overall volume of trade during the last months of 2007, as seen in Table 2. The mixed order flow (CH and HC order flow) typically contributes with about the same share to the explained variance in the euro-dollar and dollar-yen exchange rate markets. In contrast, in the euro-yen exchange rate market HC order flow makes up the dominant part of the variance share, which is consistent with our discussion of computers taking advantage of triangular arbitrage opportunities in this market.

Overall, about 15 to 35 percent of the total variation in returns can be attributed to shocks to the four order flows. However, in most currency pairs, very little of this ultimate long-run price discovery that occurs via order flow does so via the CC order flow. Similar to Table 7, we also report in Table 9 the fraction of the explained share of the return variance that can be attributed to the different order flow combinations. Comparing these to the fraction of overall volume that is due to these combinations of computers and humans, reported in Table 2, gives an idea of whether the different order flow combinations contribute proportionately to the variance in returns. It is clear that CC order flow tends to contribute disproportionately little to the long-run variance of returns, and that HH order flow often contributes disproportionately much.

7 Conclusion

Using highly-detailed high-frequency trading data for three major exchange rates over 2006 and 2007, we analyze the impact of the growth of algorithmic trading on the spot interdealer foreign exchange market. We focus on the following questions: (i) Are the algorithms underlying the computer-generated trades similar enough to result in highly correlated strategies, which some fear may cause disruptive market behavior? (ii) Does algorithmic trading increase volatility in the market, perhaps as a result of the previous concern? (iii) Do algorithmic traders improve or reduce market liquidity at times of market stress? (iv) Are human or computer traders the more "informed" traders in the market, i.e. who has the most impact on price discovery? (v) Is there evidence in this market that the liquidity providers (the "makers") and not just the liquidity "takers", are informed, and do computer makers post orders more strategically than human makers?

The first three questions directly address concerns that have been raised recently in the financial press, especially in conjunction with the current crisis, while the last two questions relate more to the empirical market microstructure literature on price discovery and order placement. Together, the analysis of these questions brings new and interesting results to the table, both from a practical and academic perspective, in an area where almost no formal research has been available.

Our empirical results provide evidence that algorithmic trades are more correlated than non-algorithmic trades, suggesting that the trading strategies used by the computer traders are less diverse than those of their human counterparts. Although this may cause some concerns regarding the disruptive potential of computer-generated trades, we do not find evidence of a positive causal relationship between the proportion of algorithmic trading in the market and the level of volatility; if anything, the evidence points towards a negative relationship, suggesting that the presence of algorithmic trading reduces volatility. As for the provision of market liquidity, we find evidence that, compared to non-algorithmic traders, algorithmic traders reduce their share of liquidity provision in the minute following major data announcements, when the probability of a price jump is very high. However, they increase their share of liquidity provision to the market over the entire hour following these announcements, which is almost always a period of elevated volatility. This empirical evidence thus suggests that computers *do* provide liquidity during periods of market stress.

To address the last two questions (price discovery and order placement), we use a high-frequency VAR framework in the tradition of Hasbrouck (1991a). We find that non-algorithmic trades account for a substantially larger share of the price movements in the euro-dollar and yen-dollar exchange rate markets than would be expected given the sizable fraction of algorithmic trades. Non-algorithmic traders are the "informed" traders in these two markets, driving price discovery. In the cross-rate, the euro-yen exchange rate market, we find that computers and humans are equally "informed," likely because of the large proportion of algorithmic trades dedicated to search for triangular arbitrage opportunities. Finally, we find that, on average, computer takers or human takers that trade on prices posted by computers do not impact prices as much as they do when they trade on prices posted by humans. One interpretation of this result is that computers place limit orders more strategically than humans do. This finding dovetails with the literature on limit order books that relaxes the common modeling assumption that liquidity providers are passive.

Overall, this study therefore provides essentially no evidence to bolster the widespread concerns about the effect of algorithmic trading on the functioning of financial markets. The lesson we take from our analysis of algorithmic trading in the interdealer foreign exchange market is that it is more how algorithmic trading is used and what it is predominantly designed to achieve that determines its impact on the market, and not primarily whether or not the order flow reaching the market is generated at high frequency by computers. In the global interdealer foreign exchange market, the rapid growth of algorithmic trading has not come at the cost of lower market quality, at least not in the data we have seen so far. Given the constant search for execution speed in financial markets and the increasing availability of algorithmic trading technology, it is likely that, absent regulatory intervention, the share of algorithmic trading across most financial markets will continue to grow. Our study offers hope that the growing presence of algorithmic trading will not have a negative impact on global financial markets.

Appendix: Robustness check of the VAR results

The impulse responses and variance decompositions in the above VAR analyses are derived under the assumption that there are no contemporaneous interactions between the different order flow components. This identifying assumption is appealing because it treats the order flow components symmetrically and ensures that the results are not driven by the ordering of the order flows in the VAR. On the other hand, it cannot be ruled out that one order flow component affects another one contemporaneously within the one-minute timespan over which each observation is sampled. If this is the case, the VAR specification that we use above would be too restrictive and the resulting impulse responses and variance decompositions would likely be biased. As discussed above, given the fairly low correlation that we observe in the VAR residuals for the different order flow equations, this does not appear to be a major concern, but since these correlations are not identical to zero it is still possible that other identification schemes would lead to different conclusions.

In this section we therefore perform a comprehensive robustness check of the VAR results by calculating upper and lower bounds on the impulse responses and variance decompositions. In particular, we consider all possible orderings of the order flows in the VARs, while imposing a triangular structure. That is, we still assume that returns are ordered last in the VAR and are thus affected contemporaneously by all order flow components, but we then consider all possible orderings for the different order flows. In the case where we split order flow into human and computer order flow, this results in just two different specifications—one where computer order flow affects human order flow contemporaneously but contemporaneous human order flow has no impact on computer order flow, and the opposite specification where human order flow affects computer order flow contemporaneously. In the case with four different order flows, there are 24 different orderings, when one allows for all possible triangular identification schemes, only imposing that returns are ordered last. From each of these specifications, we calculate impulse responses and variance decompositions. The minimum and maximum of these over all specifications are reported in Tables A1 and A2 for the two order flow case and in Tables A3 and A4 in the four order flow case.

Starting with the simpler case with order flow split up into human or computer order flow, Tables A1 and A2 confirm our conjecture that the low correlation in the VAR residuals render the VAR specification very robust to the ordering of the order flows. The min-max intervals shown in the two tables are generally very tight and all of our earlier qualitative conclusions that we draw from our preferred structural VAR specification holds also under these alternative orderings.

Turning to the VAR analysis with four separate order flow, the number of possible orderings increases dramatically to 24. This large number of possible specifications inevitably results in wider min-max intervals, even though the correlations in the VAR residuals are generally small. In order to usefully interpret these results, we check whether our main qualitative conclusions from our preferred structural specification analyzed above also holds up, in a min-max sense, under all possible orderings. Our first main result in the above analysis was that there is an initial over-reaction to CC and CH shocks and an initial under-reaction to HH shocks. As seen in Table A3, these findings are mostly supported by the min-max results as well. The only exceptions recorded are for the euro-yen cross rate, where the under-reaction to CC and CH shocks is also much weaker in the original results in Table 8. It is also evident from Table A3, Panel B, that the min-max results support the finding that a one standard deviation shock to HH has a substantially bigger impact on returns than a CC shock. In addition, Table A3, Panel B, also shows that the impact of the HH shock tends to be larger than the CH impact, and the CC impact tends to be smaller than the HC impact. Finally, the results in Table A3 also mostly support the finding that the reactions to HC order flow are greater in the euro-yen cross currency than in the two main currency pairs, although some overlap is seen for the one standard deviation shock in Panel B.

Table A4 shows the corresponding min-max results for the variance decomposition. Again, our main conclusions are mostly supported in a min-max sense. HH makes up the largest share of the explained variance in the two main currency pairs in the full sample, although in the three-month sub-sample there is some overlap between the min-max intervals for the HH order flow and the HC order flow. CC always contributes a very small share of the explained variance and HC always contributes a fairly substantial share in the cross currency.

In summary, these robustness checks show that our main VAR used for examining price discovery (equation (2)), using human and computer order flows, is not particularly sensitive to the exact identification scheme that is used. The results presented in Tables 6 and 7 thus appear to be robust to alternative orderings in the VAR. Our second VAR specification (equation (3)), which we use to analyze strategic liquidity provision, is a little more sensitive to the exact identification scheme used, but the min-max results are still overall very supportive of our main conclusions.

Table A1: Min-max impulse responses from the VAR specification with human-taker and computer-taker order flow. The table shows the minimum and maximum triangular impulse responses for returns as a result of shocks to the human-taker order flow ($HH + CH$) or computer-taker ($CC + HC$) order flow, denoted H-taker and C-taker in the table headings, respectively. In Panel A we show the return response, in basis points, to a one-billion base-currency shock to one of the order flows. In Panel B we show the return response, in basis points, to a one standard deviation shock to one of the order flows. We show the results for the full 2006-2007 sample and for the three-month sub-sample, which only uses data from September, October, and November of 2007. For each currency pair we show the short-run (immediate) response of returns; the corresponding cumulative long-run response of returns, calculated as the cumulative impact of the shock after 30 minutes; and the difference between the cumulative long-run response in returns minus the immediate response of returns, i.e., we provide the extent of over-reaction or under-reaction to an order flow shock. There are a total of 717,120 minute-by-minute observations in the full two-year sample and 89,280 observations in the three-month sub-sample.

	Full 2006-2007 sample		3-month sub-sample	
	H-taker	C-taker	H-taker	C-taker
	Panel A: One billion base-currency shock			
	USD/EUR			
Short run	[28.05, 28.06]	[26.84, 26.94]	[23.11, 23.20]	[24.89, 25.22]
Long run	[27.85, 27.87]	[32.26, 32.35]	[24.06, 24.16]	[31.04, 31.38]
Difference	[−0.20, −0.20]	[5.42, 5.42]	[0.94, 0.96]	[6.14, 6.16]
	JPY/USD			
Short run	[44.96, 46.76]	[28.92, 39.81]	[44.99, 48.02]	[33.45, 44.88]
Long run	[45.50, 47.50]	[33.21, 44.27]	[46.80, 49.54]	[28.83, 40.63]
Difference	[0.54, 0.74]	[4.29, 4.46]	[1.52, 1.81]	[−4.62, −4.25]
	JPY/EUR			
Short run	[90.18, 99.32]	[90.50, 102.71]	[109.04, 124.02]	[101.74, 115.52]
Long run	[98.30, 108.07]	[96.57, 109.85]	[116.54, 132.53]	[108.54, 123.26]
Difference	[8.12, 8.75]	[6.07, 7.14]	[7.50, 8.51]	[6.79, 7.74]
	Panel B: One standard deviation shock			
	USD/EUR			
Short run	[0.6613, 0.6616]	[0.2630, 0.2639]	[0.6023, 0.6045]	[0.3139, 0.3180]
Long run	[0.6566, 0.6570]	[0.3161, 0.3170]	[0.6269, 0.6296]	[0.3914, 0.3957]
Difference	[−0.0047, −0.0046]	[0.0531, 0.0531]	[0.0246, 0.0251]	[0.0775, 0.0777]
	JPY/USD			
Short run	[0.8370, 0.8660]	[0.2375, 0.3251]	[0.9594, 1.0158]	[0.3798, 0.5056]
Long run	[0.8470, 0.8796]	[0.2727, 0.3616]	[0.9980, 1.0480]	[0.3274, 0.4577]
Difference	[0.0100, 0.0137]	[0.0352, 0.0364]	[0.0322, 0.0386]	[−0.0524, −0.0479]
	JPY/EUR			
Short run	[0.5060, 0.5541]	[0.4318, 0.4874]	[0.6671, 0.7532]	[0.6725, 0.7581]
Long run	[0.5515, 0.6030]	[0.4608, 0.5213]	[0.7130, 0.8049]	[0.7174, 0.8089]
Difference	[0.0455, 0.0488]	[0.0289, 0.0339]	[0.0459, 0.0517]	[0.0449, 0.0508]

Table A2: Min-max variance decompositions from the VAR specification with human-taker and computer-taker order flow. The table shows the minimum and maximum triangular long-run variance decomposition of returns, expressed in percent and calculated at the 30 minute horizon. That is, the table shows the proportion of the long-run variation in returns that can be attributed to shocks to the human-taker order flow ($HH + CH$) and the computer-taker ($CC + HC$) order flow, denoted H-taker and C-taker in the table headings, respectively. For each currency pair we show the actual variance decomposition, and the proportion of the *explained* variance in returns that can be attributed to each order flow type. That is, we re-scale the variance decompositions so that they add up to 100 percent. We show results for the full 2006-2007 sample and for the three-month sub-sample, which only uses data from September, October, and November of 2007. There are a total of 717,120 minute-by-minute observations in the full two-year sample and 89,280 observations in the three-month sub-sample.

	Full 2006-2007 sample		3-month sub-sample	
	H-taker	C-taker	H-taker	C-taker
	USD/EUR			
Variance decomposition	[29.25, 29.28]	[4.71, 4.74]	[25.78, 25.96]	[7.08, 7.26]
Proportion of *explained* share	[86.04, 86.14]	[13.86, 13.96]	[78.02, 78.58]	[21.42, 21.98]
	JPY/USD			
Variance decomposition	[27.71, 29.67]	[2.31, 4.28]	[26.03, 29.19]	[4.21, 7.37]
Proportion of *explained* share	[86.63, 92.77]	[7.23, 13.37]	[77.94, 87.40]	[12.60, 22.06]
	JPY/EUR			
Variance decomposition	[10.15, 12.16]	[7.37, 9.39]	[9.94, 12.67]	[10.15, 12.88]
Proportion of *explained* share	[51.93, 62.27]	[37.73, 48.07]	[43.55, 55.53]	[44.47, 56.45]

Table A3: Min-max impulse responses from the VAR specification with all four human/computer-maker/taker order flow combinations. The table shows the minimum and maximum triangular impulse responses for returns as a result of shocks to the human-maker/human-taker order flow (HH), computer-maker/human-taker order flow (CH), human-maker/computer-taker order flow (HC), or computer-maker/computer-taker order flow (CC), denoted in obvious notation in the table headings. The results are based on estimation of equation (3), using minute-by-minute data. In Panel A we show the return response, in basis points, to a one-billion base-currency shock to one of the order flows. In Panel B we show the return response, in basis points, to a one standard deviation shock to one of the order flows. We report the results for the full 2006-2007 sample and for the three-month sub-sample, which only uses data from September, October, and November of 2007. For each currency pair we show the short-run (immediate) response of returns; the corresponding cumulative long-run response of returns, calculated as the cumulative impact of the shock after 30 minutes; and the difference between the cumulative long-run response in returns minus the immediate response of returns, i.e., we provide the extent of over-reaction or under-reaction to an order flow shock. There are a total of 717, 120 minute-by-minute observations in the full two-year sample and 89, 280 observations in the three-month sub-sample.

	Full 2006-2007 sample				3-month sub-sample			
	H-maker/ H-taker	C-maker/ H-taker	H-maker/ C-taker	C-maker/ C-taker	H-maker/ H-taker	C-maker/ H-taker	H-maker/ C-taker	C-maker/ C-taker
	Panel A: One billion base-currency shock							
	USD/EUR							
Short run	[27.33, 33.32]	[18.63, 53.71]	[15.26, 45.85]	[20.49, 52.41]	[20.21, 27.69]	[17.49, 47.57]	[17.01, 42.58]	[12.66, 38.58]
Long run	[29.89, 34.83]	[8.21, 46.56]	[22.08, 50.08]	[12.17, 47.64]	[24.08, 30.77]	[7.50, 42.68]	[25.64, 49.51]	[−4.72, 26.10]
Difference	[1.05, 2.88]	[−10.48, −7.05]	[2.69, 7.20]	[−9.13, −3.56]	[2.26, 4.55]	[−10.15, −4.71]	[4.31, 9.43]	[−17.81, −11.64]
	JPY/USD							
Short run	[43.11, 51.88]	[39.54, 95.65]	[19.41, 60.20]	[51.93, 83.64]	[41.51, 54.04]	[40.20, 94.60]	[23.08, 64.84]	[61.16, 89.94]
Long run	[46.68, 54.24]	[29.22, 89.09]	[25.02, 63.23]	[43.96, 77.69]	[46.49, 57.52]	[36.14, 90.68]	[19.96, 59.37]	[44.83, 71.65]
Difference	[2.25, 3.71]	[−10.42, −6.47]	[1.75, 5.81]	[−8.45, −4.72]	[3.48, 4.98]	[−7.41, −1.59]	[−7.05, −2.82]	[−18.43, −15.85]
	JPY/EUR							
Short run	[102.33, 123.80]	[30.23, 132.09]	[78.32, 121.12]	[99.99, 113.19]	[138.95, 178.40]	[30.01, 147.48]	[90.10, 137.21]	[86.01, 101.70]
Long run	[115.86, 137.80]	[25.52, 136.33]	[84.82, 129.72]	[91.10, 105.26]	[159.08, 198.13]	[20.15, 146.71]	[96.18, 144.85]	[86.58, 102.72]
Difference	[13.37, 14.59]	[−4.75, 4.27]	[6.14, 8.73]	[−8.89, −7.87]	[18.60, 21.43]	[−9.86, −0.77]	[4.46, 7.81]	[0.53, 1.67]
	Panel B: One standard deviation shock							
	USD/EUR							
Short run	[0.4906, 0.6497]	[0.1545, 0.4421]	[0.1276, 0.3793]	[0.0551, 0.1454]	[0.3895, 0.5839]	[0.1739, 0.4644]	[0.1823, 0.4469]	[0.0481, 0.1498]
Long run	[0.5365, 0.6791]	[0.0681, 0.3832]	[0.1847, 0.4143]	[0.0327, 0.1321]	[0.4641, 0.6489]	[0.0746, 0.4164]	[0.2747, 0.5157]	[−0.0179, 0.1014]
Difference	[0.0202, 0.0537]	[−0.0868, −0.0580]	[0.0214, 0.0609]	[−0.0246, −0.0099]	[0.0468, 0.0919]	[−0.1008, −0.0460]	[0.0432, 0.1029]	[−0.0677, −0.0452]
	JPY/USD							
Short run	[0.6273, 0.8020]	[0.2540, 0.5984]	[0.1395, 0.4108]	[0.1244, 0.2043]	[0.6484, 0.9038]	[0.3427, 0.7866]	[0.2224, 0.5924]	[0.2306, 0.3453]
Long run	[0.6793, 0.8384]	[0.1877, 0.5574]	[0.1799, 0.4315]	[0.1053, 0.1898]	[0.7261, 0.9619]	[0.3080, 0.7540]	[0.1923, 0.5424]	[0.1690, 0.2751]
Difference	[0.0347, 0.0547]	[−0.0669, −0.0405]	[0.0118, 0.0418]	[−0.0203, −0.0116]	[0.0581, 0.0777]	[−0.0576, −0.0143]	[−0.0633, −0.0272]	[−0.0708, −0.0597]
	JPY/EUR							
Short run	[0.4281, 0.5356]	[0.0847, 0.3515]	[0.3459, 0.5018]	[0.1549, 0.1755]	[0.5595, 0.7502]	[0.1084, 0.5103]	[0.5330, 0.7700]	[0.2191, 0.2590]
Long run	[0.4847, 0.5963]	[0.0715, 0.3628]	[0.3746, 0.5374]	[0.1412, 0.1632]	[0.6405, 0.8332]	[0.0728, 0.5076]	[0.5690, 0.8128]	[0.2206, 0.2616]
Difference	[0.0565, 0.0620]	[−0.0133, 0.0114]	[0.0250, 0.0362]	[−0.0138, −0.0122]	[0.0775, 0.0884]	[−0.0356, −0.0027]	[0.0244, 0.0466]	[0.0014, 0.0043]

Table A4: Min-max variance decompositions from the VAR specification with all four human/computer-maker/taker order flow combinations. The table shows the minimum and maximum triangular long-run variance decomposition of returns, expressed in percent and calculated at the 30 minute horizon. That is, the table shows the proportion of the long-run variation in returns that can be attributed to shocks to the human-maker/human-taker order flow (HH), computer-maker/human-taker order flow (CH), human-maker/computer-taker order flow (HC), and computer-maker/computer-taker order flow (CC), denoted in obvious notation in the table headings. We show the actual variance decomposition, and the proportions of the *explained* variance in returns that can be attributed to each order flow type. That is, we re-scale the variance decompositions so that they add up to 100 percent. We present results for the full 2006-2007 sample and for the three-month sub-sample, which only uses data from September, October, and November of 2007. There are a total of 717, 120 minute-by-minute observations in the full two-year sample and 89, 280 observations in the three-month sub-sample.

	Full 2006-2007 sample				3-month sub-sample			
	H-maker/ H-taker	C-maker/ H-taker	H-maker/ C-taker	C-maker/ C-taker	H-maker/ H-taker	C-maker/ H-taker	H-maker/ C-taker	C-maker/ C-taker
	USD/EUR							
Variance decomp.	[16.15, 28.28]	[1.64, 13.06]	[1.16, 9.71]	[0.21, 1.42]	[10.90, 24.29]	[2.22, 15.27]	[2.50, 14.24]	[0.23, 1.63]
Proportion	[47.42, 83.05]	[4.80, 38.36]	[3.42, 28.51]	[0.63, 4.17]	[32.50, 72.41]	[6.61, 45.52]	[7.46, 42.43]	[0.69, 4.85]
	JPY/USD							
Variance decomp.	[15.61, 25.47]	[2.57, 14.16]	[0.85, 6.77]	[0.61, 1.65]	[12.00, 23.18]	[3.34, 17.45]	[1.51, 10.04]	[1.52, 3.40]
Proportion	[48.22, 78.70]	[7.94, 43.74]	[2.64, 20.92]	[1.90, 5.10]	[35.04, 67.65]	[9.74, 50.94]	[4.40, 29.29]	[4.44, 9.91]
	JPY/EUR							
Variance decomp.	[7.29, 11.38]	[0.29, 4.89]	[4.73, 9.95]	[0.95, 1.22]	[7.02, 12.58]	[0.29, 5.83]	[6.41, 13.31]	[1.07, 1.50]
Proportion	[37.25, 58.18]	[1.49, 24.98]	[24.19, 50.87]	[4.87, 6.24]	[30.40, 54.54]	[1.26, 25.27]	[27.79, 57.67]	[4.66, 6.49]

References

[1] Albuquerque, R., and Miao, J., 2008, Advanced Information and Asset Prices, Working Paper Boston University.

[2] Andersen, T.G., and T. Bollerslev, 1998, Deutsche Mark-Dollar Volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies, *Journal of Finance* 53, 219-265.

[3] Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens, 2001. The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43–76.

[4] Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys, 2003. Modeling and Forecasting Realized Volatility, *Econometrica* 71, 579-625.

[5] Andersen, T.G., T. Bollerslev, F.X. Diebold, and C. Vega, 2003. Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange, *American Economic Review* 93, 38-62.

[6] Andersen, T.G., T. Bollerslev, F.X. Diebold, and C. Vega, 2007. Real-time price discovery in global stock, bond, and foreign exchange markets, *Journal of International Economics* 73, 251-277.

[7] Angel, J., 1992, Limit versus Market Orders, working paper, Georgetown University.

[8] Bandi, F.M., and J.R. Russel, 2006. Separating microstructure noise from volatility, *Journal of Financial Economics* 79, 655-692.

[9] Berger, D.W., A.P. Chaboud, S.V. Chernenko, E. Howorka, and J.H. Wright, 2008. Order Flow and Exchange Rate Dynamics in Electronic Brokerage System Data, *Journal of International Economics* 75, 93-109.

[10] Biais, B., Hillion, P., and Spatt, C., 1995, An empirical analysis of the limit order book and the order flow in the Paris Bourse, *Journal of Finance*, 50, 1655-1689.

[11] Bollerslev T., and J.H. Wright, 2000. Semiparametric estimation of long-memory volatility dependencies: The role of high-frequency data, *Journal of Econometrics* 98, 81-106.

[12] Bloomfield, R., M. O'Hara and G. Saar, 2005, The "Make or Take" Decision in an Electronic Market: Evidence on the Evolution of Liquidity, *Journal of Financial Economics* 75, 165–199.

[13] Chaboud, A., B. Chiquoine, E. Hjalmarsson, and M. Loretan, 2007. Frequency of Observation and the Estimation of Integrated Volatility in Deep and Liquid Markets, *Journal of Empirical Finance*, Forthcoming.

[14] Chakravarty, S., and C. Holden, 1995, An Integrated Model of Market and Limit Orders, *Journal of Financial Intermediation* 4, 213–241.

[15] Easley, D., and O'Hara, M., 1987, Price, trade size, and information in securities markets, *Journal of Financial Economics*, 19, 69-90.

[16] Evans, M., and R. Lyons, 2002. Order Flow and Exchange Rate Dynamics, *Journal of Political Economy* 110, 170-180.

[17] Evans, M., and R. Lyons, 2008. How is Macro News Transmitted to Exchange Rates?, *Journal of Financial Economics* 88, 26-50.

[18] Goettler, R., C. Parlour and U. Rajan, 2007, Microstructure effects and asset pricing, working paper, UC Berkeley.

[19] Harris, L., 1998, Optimal Dynamic Order Submission Strategies in Some Stylized Trading Problems, *Financial Markets, Institutions & Instruments* 7, 1–76.

[20] Hasbrouck, J., 1991a. Measuring the information content of stock trades, *Journal of Finance* 46, 179-207.

[21] Hasbrouck, J., 1991b. The summary informativeness of stock trades: An econometric analysis, *Review of Financial Studies* 4, 571-595.

[22] Hasbrouck, J., 1996. Order characteristics and stock price evolution: An application to program trading, *Journal of Financial Economics* 41, 129-149.

[23] He, H., and J. Wang, 1995, Differential information and dynamic behavior of stock trading volume, *Review of Financial Studies*, 8, 919-972.

[24] Hendershott, T., C.M. Jones, and A.J. Menkveld, 2007. Does algorithmic trading improve liquidity?, working paper, University of California, Berkeley.

[25] Kaniel, R., and H. Liu, 2006, So What Orders do Informed Traders Use?, *Journal of Business* 79, 1867–1913.

[26] Kumar, P., and D. Seppi, 1994, Limit and Market Orders with Optimizing Traders, working paper, Carnegie Mellon University.

[27] Madhavan, A., Richardson, M., and Roomans, M., 1997, Why do security prices change? A transaction-level analysis of NYSE stocks, *Review of Financial Studies*, 10, 1035-1064.

[28] Milgrom, P., and Stokey, N., 1982, Information, trade and common knowledge, *Journal of Economic Theory*, 26, 17-27.

[29] Parlour, Christine, and Seppi, Duane, 2008, Limit Order Markets: A Survey, *Handbook of Financial Intermediation & Banking*, ed. A.W.A. Boot and A.V. Thakor, forthcoming.

[30] Pasquariello, Paolo, and Vega, Clara, 2007, Informed and Strategic Order Flow in the Bond Markets, *Review of Financial Studies*, 20, 1975-2019.

[31] Stock, J.H., J.H. Wright, and M. Yogo, 2005. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments, *Journal of Business and Economic Statistics*, 20, 518-529.

[32] Stock, J.H., and M. Yogo, 2005. Testing for Weak Instruments in Linear IV Regression, in D.W.K. Andrews and J.H. Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge: Cambridge University Press, 80–108.

[33] Stoffman, Noah, 2007. Who Trades With Whom? Individuals, Institutions, and Returns. Unpublished Manuscript

Table 1: Summary statistics for the one-minute return and order flow data. The mean and standard deviation, as well as the first-order autocorrelation, ρ, are shown for each variable and currency pair. The returns are expressed in basis points and the order flows in millions of the base currency. The summary statistics are given for both the full 2006-2007 sample, as well as for the three-month sub-sample, which only uses observations from September, October, and November of 2007. The first two rows for each currency show the summary statistics for returns and the total market-wide order flow. The following two rows give the results for the order flows broken down into human-takers and computer-takers and the last four rows show the results for the order flow decomposed into each maker-taker pair. There are a total of $717,120$ observations in the full two-year sample and $89,280$ observations in the three-month sub sample. We show the statistical significance of the first order autocorrelation. The ***, **, and * represent significance at the 1, 5, and 10 percent level, respectively.

	Full 2006-2007 Sample			3-month sub sample		
Variable	Mean	Std. dev.	ρ	Mean	Std. dev.	ρ
			USD/EUR			
Returns	0.0030	1.2398	-0.005^{***}	0.0080	1.2057	0.007^{**}
Total order flow $(HH+CH+HC+CC)$	0.0315	25.9455	0.150^{***}	-0.0937	29.7065	0.174^{***}
H-taker $(HH+CH)$	0.0413	23.977	0.155^{***}	-0.0796	26.8096	0.189^{***}
C-taker $(HC+CC)$	-0.0099	9.9363	0.127^{***}	-0.0140	12.8900	0.115^{***}
H-maker/H-taker (HH)	0.1425	19.9614	0.177^{***}	0.0327	21.9211	0.209^{***}
C-maker/H-taker (CH)	-0.1012	8.8970	0.166^{***}	-0.1123	10.7649	0.215^{***}
H-maker/C-taker (HC)	0.0123	8.9232	0.152^{***}	0.0483	11.5856	0.150^{***}
C-maker/C-taker (CC)	-0.0222	2.7939	0.053^{***}	-0.0623	3.9477	0.072^{***}
			JPY/USD			
Returns	-0.0007	1.6038	-0.010^{***}	-0.0045	1.9110	0.007^{**}
Total order flow $(HH+CH+HC+CC)$	0.1061	20.0980	0.189^{***}	-0.3439	23.6359	0.211^{***}
H-taker $(HH+CH)$	0.0853	19.1127	0.190^{***}	-0.2088	22.0344	0.204^{***}
C-taker $(HC+CC)$	0.0209	8.3941	0.170^{***}	-0.1351	11.5877	0.158^{***}
H-maker/H-taker (HH)	0.1037	15.9972	0.209^{***}	-0.1203	17.4612	0.226^{***}
C-maker/H-taker (CH)	-0.0184	6.9030	0.172^{***}	-0.0885	9.1773	0.162^{***}
H-maker/C-taker (HC)	0.0198	7.5686	0.198^{***}	-0.0901	10.1673	0.191^{***}
C-maker/C-taker (CC)	0.0011	2.4556	0.032^{***}	-0.045	3.8751	0.026^{***}
			JPY/EUR			
Returns	0.0024	1.5976	-0.053^{***}	0.0036	2.1398	-0.017^{***}
Total order flow $(HH+CH+HC+CC)$	-0.0648	7.0941	0.152^{***}	-0.1574	8.5978	0.147^{***}
H-taker $(HH+CH)$	-0.0497	5.7006	0.150^{***}	-0.1216	6.2074	0.125^{***}
C-taker $(HC+CC)$	-0.0151	4.8409	0.146^{***}	-0.0358	6.7000	0.131^{***}
H-maker/H-taker (HH)	-0.0172	4.4203	0.159^{***}	-0.0600	4.3106	0.157^{***}
C-maker/H-taker (CH)	-0.0325	2.8912	0.129^{***}	-0.0616	3.7197	0.092^{***}
H-maker/C-taker (HC)	-0.0095	4.5331	0.173^{***}	-0.0264	6.0968	0.161^{***}
C-maker/C-taker (CC)	-0.0056	1.5558	0.023^{***}	-0.0095	2.5621	-0.001

Table 2: Summary statistics for the fractions of trade volume attributable to different trader combinations. The table shows the fraction of the total volume of trade that is attributable to different combinations of makers and takers. Results for the full 2006-2007 sample as well as for the three-month sub-sample, which only uses data from September, October, and November of 2007, are shown. We show the average of the daily fractions, calculated by summing up across all minutes within a day, and the standard deviations of those daily fractions. For each currency, the first row shows the fraction of the total volume of trade where a computer was involved on at least one side of the trade (i.e. as a maker or a taker). The second row shows the fraction of the total volume where a human acted as a taker, the third row shows the fraction of the total volume where a computer acted as a taker, and the following four rows shows the fractions broken down by each maker-taker pair.

	Full 2006-2007 Sample		3-month sub sample	
Variable	Mean	Std. dev.	Mean	Std. dev.
	USD/EUR			
C-participation ($Vol\,(CH+HC+CC)$)	0.4163	0.1135	0.5386	0.0355
H-taker ($Vol\,(CH+HH)$)	0.7810	0.0791	0.6864	0.0331
C-taker ($Vol\,(HC+CC)$)	0.2190	0.0791	0.3136	0.0331
H-maker/H-taker ($Vol\,(HH)$)	0.5837	0.1135	0.4614	0.0355
C-maker/H-taker ($Vol\,(CH)$)	0.1973	0.0398	0.2251	0.0144
H-maker/C-taker ($Vol\,(HC)$)	0.1710	0.0514	0.2304	0.0205
C-maker/C-taker ($Vol\,(CC)$)	0.0480	0.0290	0.0831	0.0150
	JPY/USD			
C-participation ($Vol\,(CH+HC+CC)$)	0.4242	0.1065	0.5652	0.0364
H-taker ($Vol\,(CH+HH)$)	0.7585	0.0805	0.6461	0.0311
C-taker ($Vol\,(HC+CC)$)	0.2415	0.0805	0.3539	0.0311
H-maker/H-taker ($Vol\,(HH)$)	0.5758	0.1065	0.4348	0.0364
C-maker/H-taker ($Vol\,(CH)$)	0.1827	0.0304	0.2114	0.0126
H-maker/C-taker ($Vol\,(HC)$)	0.1860	0.0498	0.2486	0.0154
C-maker/C-taker ($Vol\,(CC)$)	0.0555	0.0321	0.1052	0.0193
	JPY/EUR			
C-involved ($Vol\,(CH+HC+CC)$)	0.6186	0.1154	0.7907	0.0410
H-taker ($Vol\,(CH+HH)$)	0.5557	0.1018	0.4037	0.0467
C-taker ($Vol\,(HC+CC)$)	0.4443	0.1018	0.5963	0.0467
H-maker/H-taker ($Vol\,(HH)$)	0.3814	0.1154	0.2093	0.0410
C-maker/H-taker ($Vol\,(CH)$)	0.1743	0.0360	0.1944	0.0164
H-maker/C-taker ($Vol\,(HC)$)	0.3337	0.0473	0.3734	0.0193
C-maker/C-taker ($Vol\,(CC)$)	0.1106	0.0673	0.2229	0.0464

Table 3: Estimates of the ratio $R = RC/RH$. The table reports the mean estimates of the ratio $R = RC/RH$, where $RC = Vol(HC)/Vol(CC)$ and $RH = Vol(HH)/Vol(CH)$. $Vol(HH)$ is the daily trading volume between human-makers and human-takers, $Vol(HC)$ is the daily trading volume between human-makers and computer-takers, $Vol(CH)$ is the daily trading volume between computer-makers and human-takers, and $Vol(CC)$ is the daily trading volume between computer-makers and computer-takers. We report the mean of the daily ratio R and the standard errors are shown in parantheses below the estimate. We also show the number of days that had a ratio that was less than one. We report the results for the full 2006-2007 sample and the three-month sub-sample, which only uses data from September, October, and November of 2007. The ***, **, and * represent a statistically significant deviation from one at the 1, 5, and 10 percent level, respectively.

	Full 2006-2007 sample	3-month sub sample
	USD/EUR	
Mean of daily $R = RC/RH$	1.4463^{***}	1.3721^{***}
Standard Error	(0.0063)	(0.0099)
No. of days with $R < 1$	0	0
No. of obs	498	62
	JPY/USD	
Mean of daily $R = RC/RH$	1.2619^{***}	1.1719^{***}
Standard Error	(0.0074)	(0.0142)
No. of days with $R < 1$	15	4
No. of obs	498	62
	JPY/EUR	
Mean of daily $R = RC/RH$	1.6886^{***}	1.6242^{***}
Standard Error	(0.0154)	(0.0250)
No. of days with $R < 1$	4	0
No. of obs	498	62

Table 4: Regressions of realized volatility on the fraction of algorithmic trading. The table shows the results from estimating the relationship between daily realized volatility and the fraction of algorithmic trading, using daily data from 2006 and 2007. Robust standard errors are given in parentheses below the coefficient estimates. The left hand side of the table shows the results with a quarterly time trend included in the regressions and the right hand side of the table shows the results with year-quarter time dummies (i.e., eight time dummies, one for each quarter in the two years of data) included in the regressions. Panels A and B show the results when the fraction of algorithmic trading is measured as the fraction of the total trade volume that has a computer involved on at least one side of the trade (i.e. as a maker or a taker). Panels C and D show the results when only the fraction of volume with computer taking is used. In addition to the fraction of algorithmic trading and the control(s) for secular trends, 22 lags of volatility are also included in every specification. In all cases, only the coefficient on the fraction of algorithmic trading is displayed. Panels A and C show the results from a standard OLS estimation, along with the R^2. Panels B and D show the results from the IV specification estimated with Limited Information Maximum Likelihood (LIML). In Panels B and D, the Stock and Yogo (2005) $F-$test of weak instruments are also shown. The critical values for Stock and Yogo's (2005) F-test are designed such that they indicate a maximal actual size for a nominal sized five percent test on the coefficient in the LIML estimation. Thus, in order for the actual size of the LIML test to be no greater than 10% (15%), the F-statistic should exceed 8.68 (5.33). There are a total of 498 daily observations in the data. The ***, **, and * represent significance at the 1, 5, and 10 percent level, respectively.

	With quarterly time trend			With year-quarter time dummies		
	USD/EUR	JPY/USD	JPY/EUR	USD/EUR	JPY/USD	JPY/EUR
	Panel A. Fraction of volume with any computer participation, OLS estimation					
Coeff. on AT	0.0029	0.0018	0.0034^{***}	0.0078^{***}	−0.0030	0.0065^{***}
	(0.0024)	(0.0021)	(0.0012)	(0.0027)	(0.0024)	(0.0016)
R^2 (%)	53.44%	61.13%	71.90%	56.73%	62.57%	73.33%
	Panel B. Fraction of volume with any computer participation, IV estimation					
Coeff. on AT	-0.0121^{*}	-0.0186^{**}	−0.0022	−0.0078	−0.0101	−0.0128
	(0.0062)	(0.0089)	(0.0039)	(0.0061)	(0.0069)	(0.0175)
F-Stat	29.58	19.46	32.18	38.17	20.89	2.25
	Panel C. Fraction of volume with computer-taking, OLS estimation					
Coeff. on AT	0.0037	−0.0027	0.0015	0.0094^{**}	−0.0034	0.0032^{**}
	(0.0036)	(0.0024)	(0.0012)	(0.0038)	(0.0027)	(0.0016)
R^2 (%)	53.39%	61.17%	71.56%	56.43%	62.55%	72.66%
	Panel D. Fraction of volume with computer-taking, IV estimation					
Coeff. on AT	-0.0160^{**}	-0.0215^{**}	−0.0007	−0.0072	−0.0122	−0.0182
	(0.0080)	(0.0109)	(0.0028)	(0.0070)	(0.0082)	(0.0291)
F-Stat	39.99	17.63	64.81	55.45	21.20	1.04

Table 5: We report the mean ratio of the exchange rate volatility (Panel A) and liquidity provision by humans and by computers (Panel B) estimated during announcement days relative to that estimated during non-announcement days. The one-hour measure is estimated using observations from 8:25 am to 9:24 am ET and the one-minute measure is estimated using 8:30 am to 8:31 am ET observations. Announcement days are defined as nonfarm payroll announcement days and non-announcement days are defined as 10 business days before and after the nonfarm payroll announcement. In each panel, we report the chi-squared and p-value of the Wald test that the ratio is equal to 1. In Panel C we report the chi-squared and p-value of the Wald test that the liquidity provision of humans during announcement days relative to non-announcement days is similar to the liquidity provision of computers. The statistics are estimated using data in the full sample from 2006 to 2007 and there are 23 observations (April 6, 2007 nonfarm payroll announcement is missing because it falls on Good Friday, when trading in the foreign exchange market is limited). Human liquidity provision, LH, is defined as the sum of human-maker/human-taker volume plus human-maker/human-taker volume divided by total volume. Computer liquidity provision, LC, is defined as the sum of computer-maker/computer-taker volume plus computer-maker/human-taker volume divided by total volume. The ***, **, and * represent significance at the 1, 5, and 10 percent level, respectively.

	USD/EUR		JPY/USD		JPY/EUR	
	Hour	Minute	Hour	Minute	Hour	Minute
	Panel A: Volatility					
$\frac{\sigma_a}{\sigma_n}$	6.236^{***}	21.704^{***}	5.595^{***}	24.812^{***}	3.697^{***}	14.403^{**}
χ^2 $(H_0 : \sigma_a = \sigma_n)$	69.86	18.76	33.34	15.45	19.37	5.96
p-value	0.0000	0.0003	0.0000	0.0008	0.0002	0.0235
	Panel B: Liquidity Provision					
Liquidity provision by humans, $\frac{LH_a}{LH_n}$	0.964^{***}	1.062^{***}	1.023	1.183^{***}	0.888^{***}	0.980
Liquidity provision by computers, $\frac{LC_a}{LC_n}$	1.132^{***}	0.871^{***}	0.974	0.652^{***}	1.227^{***}	1.151
χ^2 $(H_0 : LH_a = LH_n or LC_a = LC_n)$	16.56	9.04	2.71	31.91	25.19	0.5
p-value	0.0005	0.0067	0.1143	0	0.0001	0.487
	Panel C: Comparison of Liquidity Provision between Humans and Computers					
$\frac{LH_a}{LH_n} - \frac{LC_a}{LC_n}$	-0.168^{***}	0.191^{**}	0.049	0.532^{***}	-0.339^{***}	-0.171
$\chi^2 \left(H_0 : \frac{LH_a}{LH_n} = \frac{LC_a}{LC_n}\right)$	19.24	5.91	1.50	36.07	25.21	0.66
p-value	0.0003	0.0241	0.2339	0.0000	0.0001	0.4245

Table 6: Impulse responses from the VAR specification with human-taker and computer-taker order flow. The table shows the impulse responses for returns as a result of shocks to the human-taker order flow $(HH + CH)$ or computer-taker $(CC + HC)$ order flow, denoted H-taker and C-taker in the table headings, respectively. The results are based on estimation of equation (2), using minute-by-minute data. In Panel A we show the return response, in basis points, to a one-billion base-currency shock to one of the order flows. In Panel B we show the return response, in basis points, to a one standard deviation shock to one of the order flows. We show the results for the full 2006-2007 sample and for the three-month sub-sample, which only uses data from September, October, and November of 2007. For each currency pair we show the short-run (immediate) response of returns; the corresponding cumulative long-run response of returns, calculated as the cumulative impact of the shock after 30 minutes; and the difference between the cumulative long-run response in returns minus the immediate response of returns, i.e., we provide the extent of over-reaction or under-reaction to an order flow shock. There are a total of $717,120$ minute-by-minute observations in the full two-year sample and $89,280$ observations in the three-month sub-sample. We show in parenthesis the standard errors of the difference between the short-run and long-run response. These standard errors are calculated by bootstrapping, using 200 repetitions.

	Full 2006-2007 sample		3-month sub-sample	
	H-taker	C-taker	H-taker	C-taker
	Panel A: One billion base-currency shock			
	USD/EUR			
Short run	28.06	26.94	23.20	25.22
Long run	27.87	32.35	24.16	31.38
Difference	−0.20	5.42	0.96	6.16
	(0.29)	(0.67)	(0.72)	(1.36)
	JPY/USD			
Short run	46.77	39.81	48.02	44.89
Long run	47.50	44.27	49.54	40.63
Difference	0.74	4.46	1.52	−4.26
	(0.48)	(1.08)	(1.36)	(2.35)
	JPY/EUR			
Short run	99.32	102.71	124.02	115.52
Long run	108.07	109.85	132.53	123.26
Difference	8.75	7.14	8.51	7.74
	(1.50)	(1.67)	(4.79)	(4.76)
	Panel B: One standard deviation shock			
	USD/EUR			
Short run	0.6617	0.2639	0.6045	0.3181
Long run	0.6570	0.3170	0.6296	0.3957
Difference	−0.0046	0.0531	0.0251	0.0777
	(0.0068)	(0.0065)	(0.0189)	(0.0172)
	JPY/USD			
Short run	0.8706	0.3269	1.0241	0.5098
Long run	0.8843	0.3635	1.0565	0.4614
Difference	0.0137	0.0366	0.0324	−0.0483
	(0.0090)	(0.0089)	(0.0289)	(0.0267)
	JPY/EUR			
Short run	0.5572	0.4901	0.7587	0.7636
Long run	0.6063	0.5242	0.8108	0.8148
Difference	0.0491	0.0341	0.0520	0.0512
	(0.0085)	(0.0080)	(0.0294)	(0.0314)

Table 7: Variance decompositions from the VAR specification with human-taker and computer-taker order flow. The table provides the long-run variance decomposition of returns, expressed in percent and calculated at the 30 minute horizon, based on estimation of equation (2), using minute-by-minute data. That is, the table shows the proportion of the long-run variation in returns that can be attributed to shocks to the human-taker order flow ($HH + CH$) and the computer-taker ($CC + HC$) order flow, denoted H-taker and C-taker in the table headings, respectively. For each currency pair we show the actual variance decomposition, and the proportion of the *explained* variance in returns that can be attributed to each order flow type. That is, we re-scale the variance decompositions so that they add up to 100 percent. We show results for the full 2006-2007 sample and for the three-month sub-sample, which only uses data from September, October, and November of 2007. There are a total of 717,120 minute-by-minute observations in the full two-year sample and 89,280 observations in the three-month sub-sample. We show in parenthesis the standard errors calculated by bootstrapping, using 200 repetitions.

	Full 2006-2007 sample		3-month sub-sample	
	H-taker	C-taker	H-taker	C-taker
	USD/EUR			
Variance decomposition	29.27	4.74	25.92	7.25
	(0.95)	(0.19)	(0.79)	(0.42)
Proportion of *explained* share	86.06	13.94	78.14	21.86
	(2.79)	(0.56)	(2.38)	(1.27)
	JPY/USD			
Variance decomposition	29.31	4.22	28.59	7.22
	(0.35)	(0.11)	(0.50)	(0.33)
Proportion of *explained* share	87.41	12.59	79.84	20.16
	(1.04)	(0.33)	(1.40)	(0.92)
	JPY/EUR			
Variance decomposition	12.03	9.28	12.47	12.67
	(0.21)	(0.20)	(0.38)	(0.38)
Proportion of *explained* share	56.45	43.55	49.60	50.40
	(0.99)	(0.94)	(1.51)	(1.51)

Table 8: Impulse responses from the VAR specification with all four human/computer-maker/taker order flow combinations. The table shows the impulse responses for returns as a result of shocks to the human-maker/human-taker order flow (HH), computer-maker/human-taker order flow (CH), human-maker/computer-taker order flow (HC), or computer-maker/computer-taker order flow (CC), denoted in obvious notation in the table headings. The results are based on estimation of equation (3), using minute-by-minute data. In Panel A we show the return response, in basis points, to a one-billion base-currency shock to one of the order flows. In Panel B we show the return response, in basis points, to a one standard deviation shock to one of the order flows. We report the results for the full 2006-2007 sample and for the three-month sub-sample, which only uses data from September, October, and November of 2007. For each currency pair we show the short-run (immediate) response of returns; the corresponding cumulative long-run response of returns, calculated as the cumulative impact of the shock after 30 minutes; and the difference between the cumulative long-run response in returns minus the immediate response of returns, i.e., we provide the extent of over-reaction or under-reaction to an order flow shock. There are a total of 717,120 minute-by-minute observations in the full two-year sample and 89,280 observations in the three-month sub-sample. We show in parenthesis the standard errors of the difference between the short-run and the long-run response. These standard errors are calculated by bootstrapping, using 200 repetitions.

	Full 2006-2007 sample				3-month sub-sample			
	H-maker/ H-taker	C-maker/ H-taker	H-maker/ C-taker	C-maker/ C-taker	H-maker/ H-taker	C-maker/ H-taker	H-maker/ C-taker	C-maker/ C-taker
	Panel A: One billion base-currency shock							
	USD/EUR							
Short run	27.64	29.66	26.57	32.19	20.58	30.94	28.94	21.74
Long run	30.13	20.47	29.89	24.92	24.18	23.35	34.64	5.94
Difference	2.49	−9.19	3.32	−7.26	3.60	−7.59	5.70	−15.80
	(0.35)	(0.88)	(0.83)	(2.42)	(0.97)	(2.03)	(1.79)	(4.64)
	JPY/USD							
Short run	43.48	58.94	40.34	61.57	41.96	64.63	46.08	67.65
Long run	47.01	49.53	42.61	54.37	46.83	57.24	40.33	51.81
Difference	3.53	−9.41	2.27	−7.20	4.87	−7.39	−5.75	−15.85
	(0.59)	(1.57)	(1.30)	(3.38)	(1.62)	(3.43)	(2.89)	(7.39)
	JPY/EUR							
Short run	102.61	92.16	100.91	102.04	139.33	103.92	114.01	94.47
Long run	116.12	91.24	107.18	93.41	159.46	96.85	118.47	95.20
Difference	13.51	−0.92	6.27	−8.63	20.13	−7.07	4.46	0.74
	(1.98)	(3.18)	(1.94)	(4.98)	(7.25)	(9.35)	(5.78)	(10.70)
	Panel B: One standard deviation shock							
	USD/EUR							
Short run	0.5389	0.2575	0.2318	0.0893	0.4342	0.3211	0.3228	0.0845
Long run	0.5875	0.1777	0.2608	0.0692	0.5101	0.2424	0.3864	0.0231
Difference	0.0486	−0.0798	0.0290	−0.0202	0.0760	−0.0788	0.0636	−0.0614
	(0.0069)	(0.0076)	(0.0072)	(0.0067)	(0.0203)	(0.0211)	(0.0200)	(0.0180)
	JPY/USD							
Short run	0.6721	0.3968	0.2962	0.1506	0.7019	0.5801	0.4544	0.2607
Long run	0.7267	0.3334	0.3129	0.1330	0.7834	0.5137	0.3976	0.1997
Difference	0.0546	−0.0634	0.0167	−0.0176	0.0815	−0.0663	−0.0567	−0.0611
	(0.0091)	(0.0106)	(0.0096)	(0.0083)	(0.0274)	(0.0307)	(0.0284)	(0.0284)
	JPY/EUR							
Short run	0.4440	0.2629	0.4481	0.1583	0.5859	0.3829	0.6809	0.2409
Long run	0.5024	0.2603	0.4760	0.1449	0.6706	0.3568	0.7076	0.2428
Difference	0.0584	−0.0026	0.0279	−0.0134	0.0847	−0.0260	0.0266	0.0019
	(0.0086)	(0.0091)	(0.0086)	(0.0077)	(0.0306)	(0.0344)	(0.0345)	(0.0273)

Table 9: Variance decompositions from the VAR specification with all four human/computer-maker/taker order flow combinations. The table provides the long-run variance decomposition of returns, expressed in percent and calculated at the 30 minute horizon, based on estimation of equation (3), using minute-by-minute data. That is, the table shows the proportion of the long-run variation in returns that can be attributed to shocks to the human-maker/human-taker order flow (HH), computer-maker/human-taker order flow (CH), human-maker/computer-taker order flow (HC), and computer-maker/computer-taker order flow (CC), denoted in obvious notation in the table headings. We show the actual variance decomposition, and the proportions of the *explained* variance in returns that can be attributed to each order flow type. That is, we re-scale the variance decompositions so that they add up to 100 percent. We present results for the full 2006-2007 sample and for the three-month sub-sample, which only uses data from September, October, and November of 2007. There are a total of 717, 120 minute-by-minute observations in the full two-year sample and 89, 280 observations in the three-month sub-sample. We show in parenthesis the standard errors, which are calculated by bootstrapping, using 200 repetitions.

	Full 2006-2007 sample				3-month sub-sample			
	H-maker/ H-taker	C-maker/ H-taker	H-maker/ C-taker	C-maker/ C-taker	H-maker/ H-taker	C-maker/ H-taker	H-maker/ C-taker	C-maker/ C-taker
	USD/EUR							
Variance decomp.	20.71	4.73	3.89	0.58	14.19	7.68	7.86	0.59
	(0.89)	(0.24)	(0.21)	(0.04)	(0.75)	(0.48)	(0.43)	(0.09)
Proportion	69.24	15.81	13.01	1.94	46.80	25.33	25.92	1.95
	(2.98)	(0.80)	(0.70)	(0.13)	(2.47)	(1.58)	(1.42)	(0.30)
	JPY/USD							
Variance decomp.	18.62	6.48	3.70	0.93	14.47	9.78	6.12	2.00
	(0.33)	(0.15)	(0.11)	(0.04)	(0.46)	(0.41)	(0.31)	(0.13)
Proportion	62.63	21.80	12.45	3.13	44.70	30.21	18.91	6.18
	(1.11)	(0.50)	(0.37)	(0.13)	(1.42)	(1.27)	(0.96)	(0.40)
	JPY/EUR							
Variance decomp.	7.84	2.74	7.94	0.99	7.72	3.32	10.47	1.30
	(0.16)	(0.12)	(0.19)	(0.06)	(0.29)	(0.20)	(0.42)	(0.11)
Proportion	40.18	14.04	40.70	5.07	33.84	14.56	45.90	5.70
	(0.82)	(0.61)	(0.97)	(0.31)	(1.27)	(0.88)	(1.84)	(0.48)

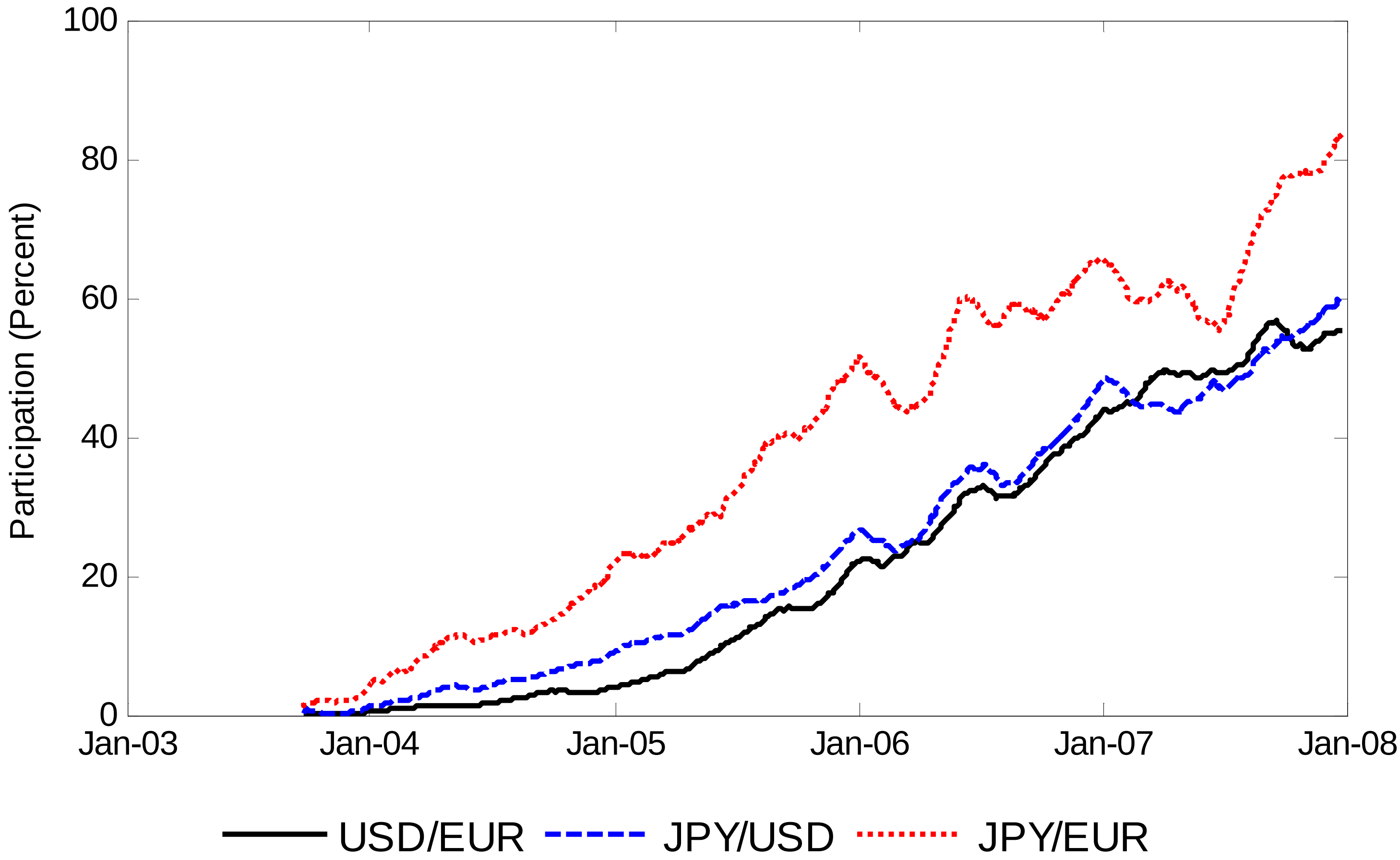


Figure 1: 50-day moving averages of participation rates of algorithmic traders

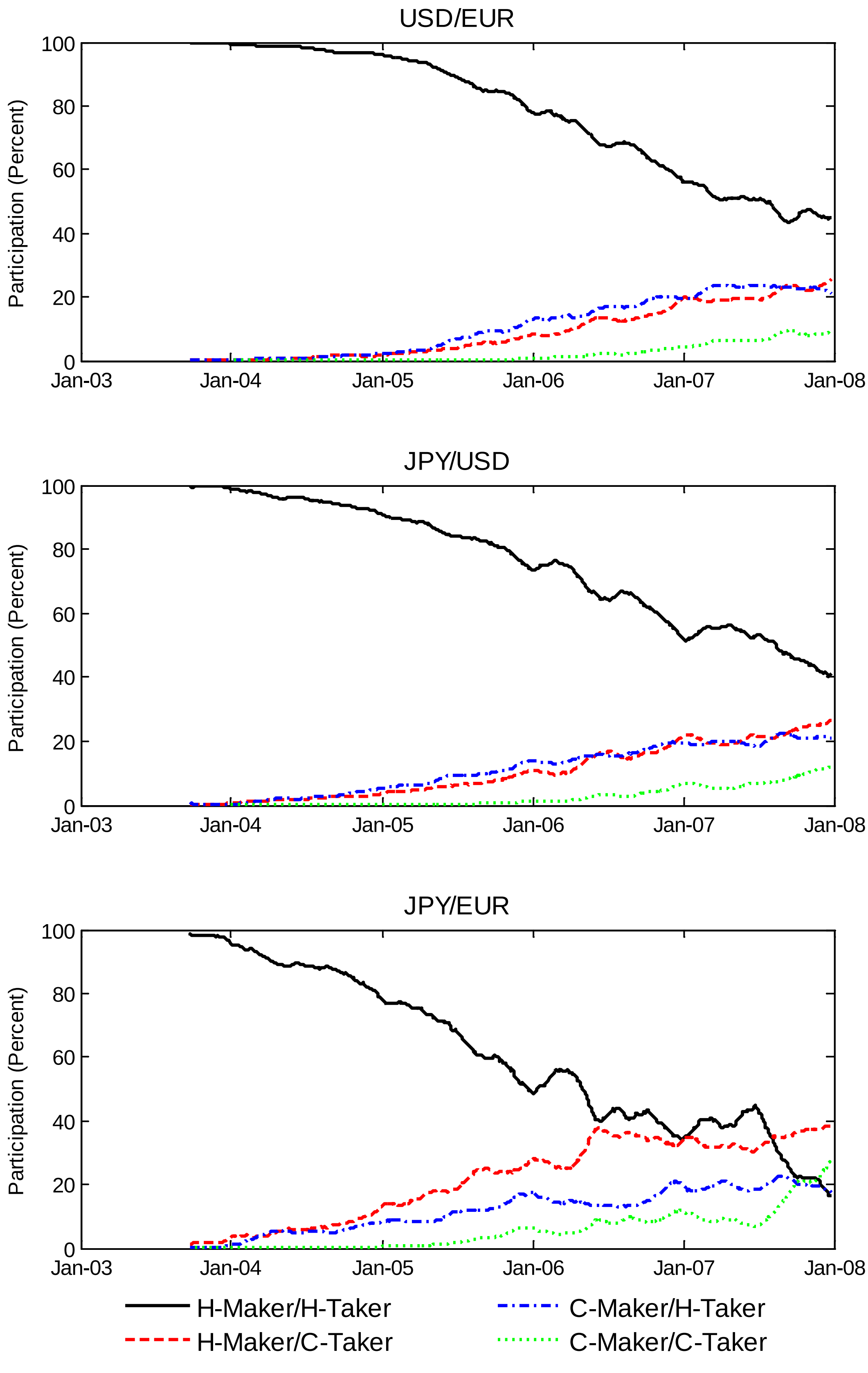
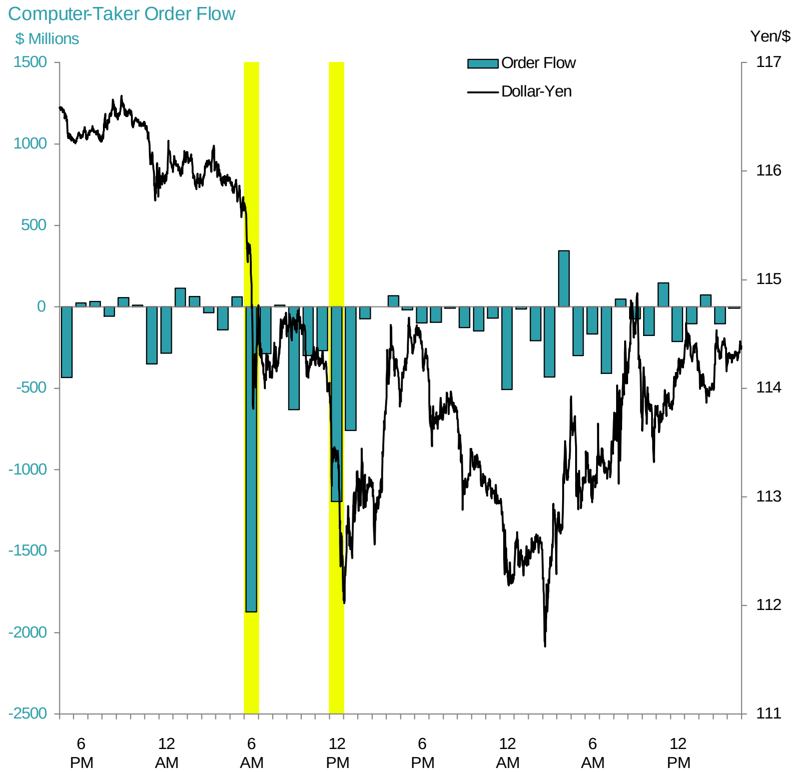


Figure 2: 50-day moving averages of participation rates broken down into four maker-taker pairs



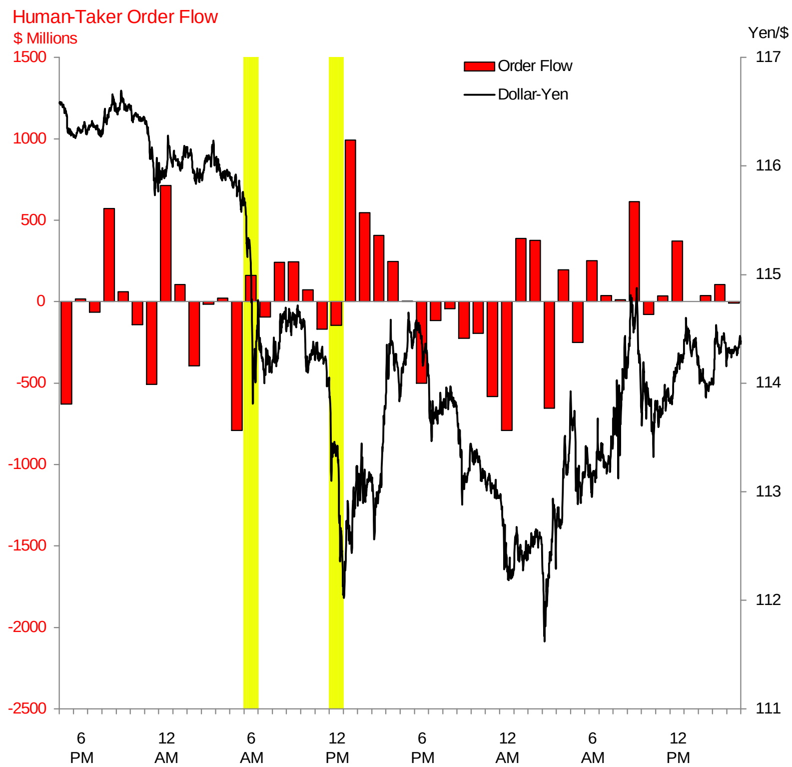


Figure 3: Dollar-Yen Market on August 16, 2007

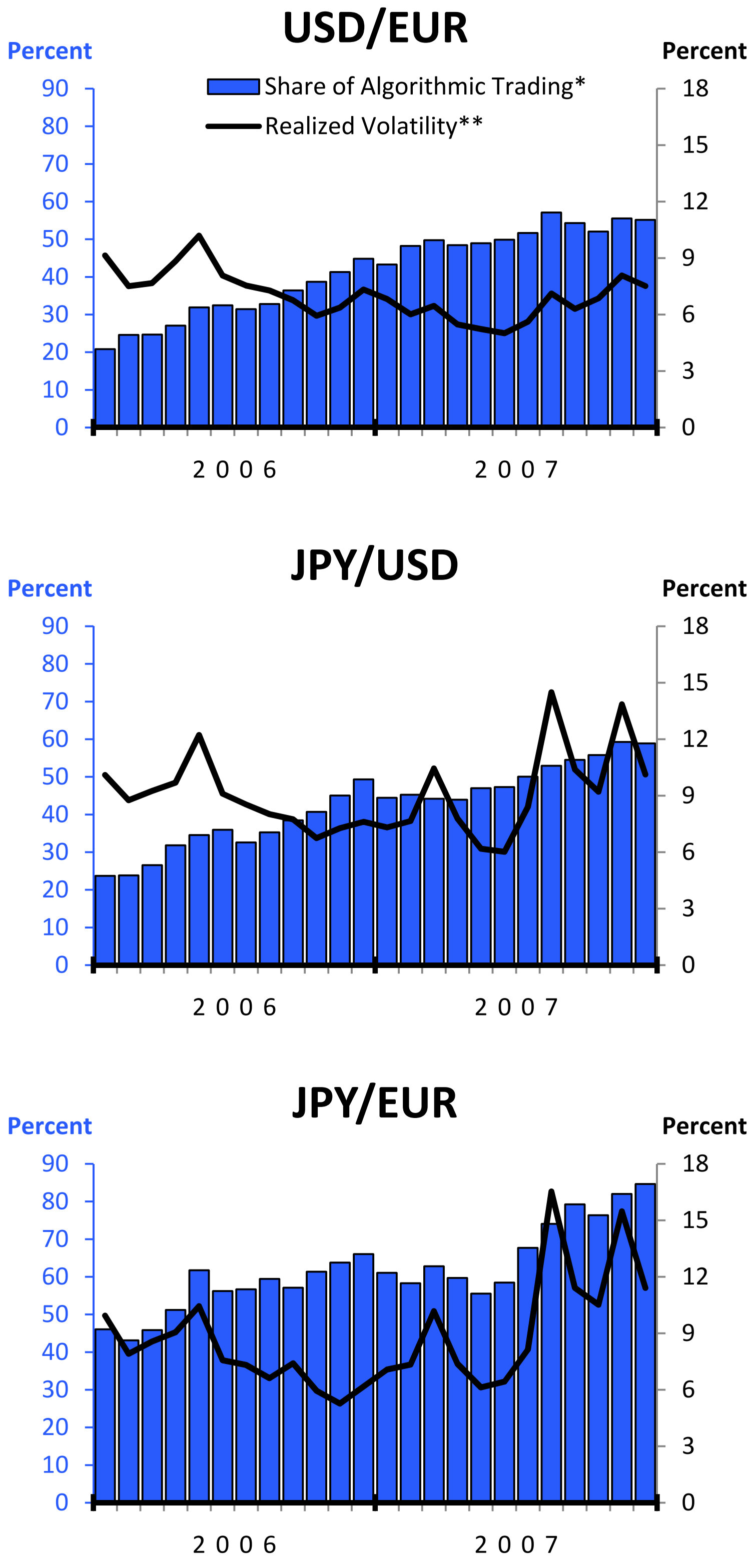


Figure 4: Volatility and Algorithmic Market Participation

[*]Daily realized volatility is based on 1-minute returns. We show monthly observations
[**]Share of algorithmic trading is at a monthly frequency

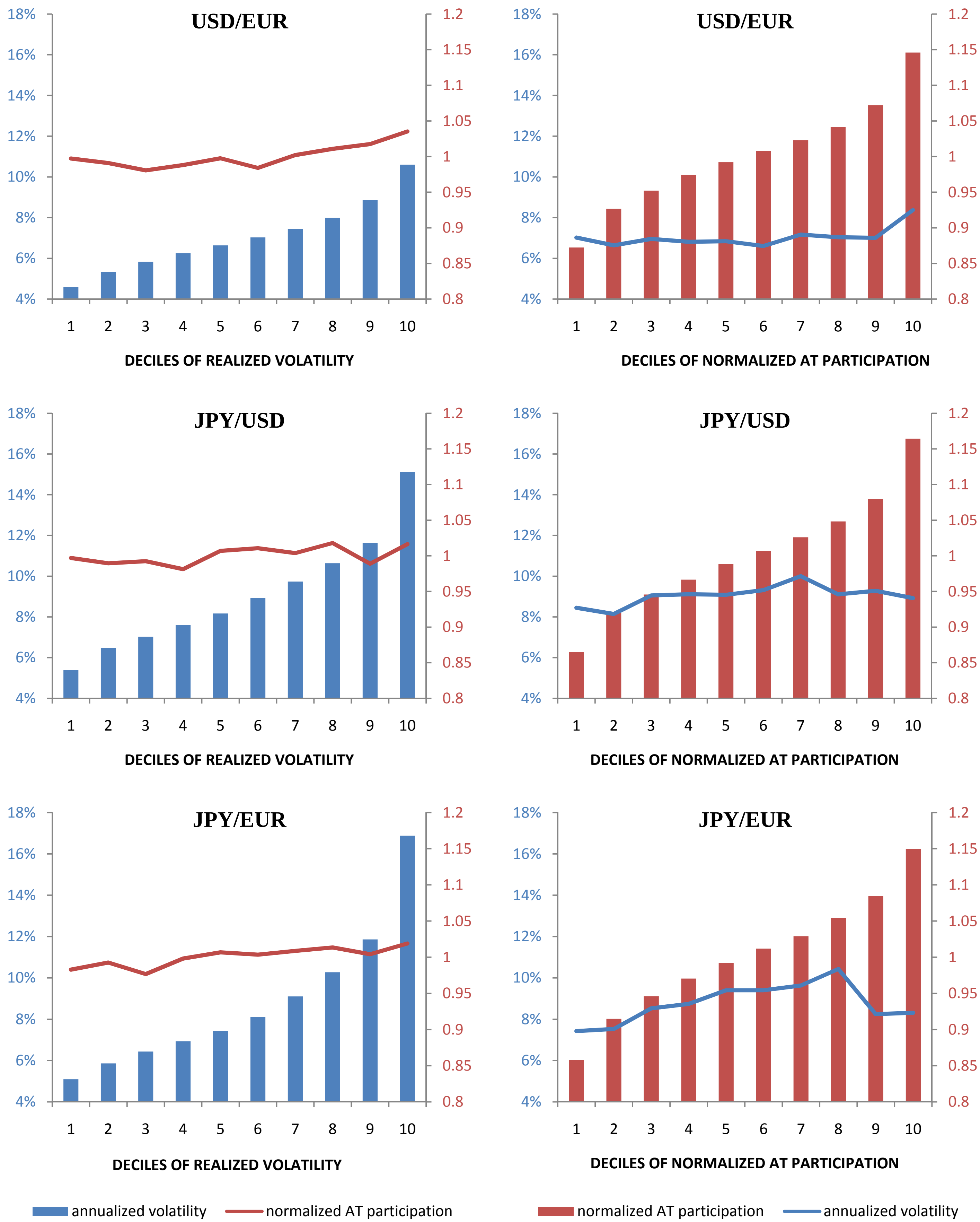


Figure 5: Deciles of Realized Volatility and AT Participation

TAB 8

High Frequency Traders and Asset Prices *

Jakša Cvitanić†
Andrei Kirilenko‡

March 11, 2010

Abstract

Do high frequency traders affect transaction prices? In this paper we derive distributions of transaction prices in limit order markets populated by low frequency traders (humans) before and after the entrance of a high frequency trader (machine). We find that the presence of a machine is likely to change the average transaction price, even in the absence of new information. We also find that in a market with a high frequency trader, the distribution of transaction prices has more mass around the center and thinner far tails. With a machine, mean intertrade duration decreases in proportion to the increase in the ratio of the human order arrival rates with and without the presence of the machine; trading volume goes up by the same rate. We show that the machine makes positive expected profits by "sniping" out human orders somewhat away from the front of the book. This explains the shape of the transaction price density. In fact, we show that in a special case, the faster humans submit and vary their orders, the more profits the machine makes.

Keywords: high-frequency trading, electronic trading, asset prices, limit orders

JEL Classification: D4, G1

---

*The research of J. Cvitanić was supported in part by NSF grant DMS 06-31298. We thank Doug Adams for useful comments on the paper. The views presented here are our own and do not represent the views of the Commodity Futures Trading Commission, its Commissioners or staff. All errors are our own.

†Caltech, Humanities and Social Sciences, M/C 228-77, 1200 E. California Blvd. Pasadena, CA 91125. Ph: (626) 395-1784. E-mail: cvitanic@hss.caltech.edu.

‡Commodity Futures Trading Commission, 1155 21st Street, NW, Washington, DC 20581. Ph: (202) 418-5587. E-mail: akirilenko@cftc.gov.

1 Introduction

High frequency trading typically refers to trading activity that employs extremely fast automated programs for generating, routing, canceling, and executing orders in electronic markets. High frequency traders submit and cancel a massive number of orders and execute a large number of trades, trade in and out of positions very quickly, and finish each trading day without a significant open position. High frequency trading is estimated to account for at least half of the trading volume on equity and derivatives exchanges.

High frequency traders are very fast, but what valuable service do they provide to the markets? Do they make prices more informative? Do they increase market liquidity? How do they make money?

In this paper we study the distribution of transaction prices generated in an electronic limit order market populated by orders from high frequency traders (machines) and low frequency traders (humans). We focus on the period between two human transactions - a very short period of time in a liquid market. We posit that during such a short horizon, the impact of changes in the fundamentals is negligible. Therefore, we model the incoming human buy order prices and sell order prices during the period as two *iid* sequences, arriving according to exogenous Poisson processes. For tractability, we assume that the submitted orders are of unit size and at infinitely divisible prices.[1] We justify this simplification as being appropriate for inter-trade intervals of relative homogeneity, in which the demand of all the traders on the buy side is approximately the same, and also close to the quantity that the individual traders on the sell side are willing to supply in a single trade.

Machines are assumed to be strategic uninformed liquidity providers. They have only one advantage over the humans - the speed with which they can submit or cancel their orders. Because of this advantage, machines dominate the trading within each period by undercutting slow humans at the front of the book. This is only one of the strategies used by actual high-frequency traders in real markets, and the only one we focus on.[2] In the language of the industry, machines aim to "pick-off" or "snipe out" incoming human orders. However, we assume that machines do not carry their submitted orders across time, for the fear of being picked off themselves. Thus, we assume that machines submit deterministic orders that get immediately canceled if not executed, and then get resubmitted again. With these actions, they shape the front of the limit order book.

When we model the optimization of the machine during the intra-trade period, we assume that it knows the process that governs the arrivals of human orders, the distribution of incoming human limit orders, and the values of existing orders in the book. In reality, the machine needs to estimate these quantities by "pinging the book" - sending quick trial orders and canceling them immediately. We do not model the estimation procedure, but assume it has been done before the beginning of the interval.

Our findings are as follows. First, we derive formulas for the distributions of transaction

[1] In the actual limit order book environment, traders submit orders of different quantities at discrete price intervals - ticks. At each tick, quantities get stacked up in accordance with a priority rule, e.g., time priority or order size and then time priority. Our idealized model, with the order prices coming from a continuous distribution and for one order only, can be thought of as taking the actual orders for multiple units stacked up at each tick and "spreading" them between ticks.

[2] Other known high frequency trading strategies include (i) the collection of rebates offered by exchanges for liquidity provision, (ii) cross-market arbitrage, and (iii) "spoofing" - triggering other traders to act.

prices and transaction times for a given intra-trade period both with and without the machines. We find that the presence of a machine is likely to change the average transaction price, even in the absence of new information. We also find that in the presence of a machine, the shape of the transaction price density remains the same in the middle, between the bid and the ask of the machine, the far tails of the density get thinner, while the parts of the tails closer to the bid and the ask of the machine get fatter. In the presence of the machine, mean intertrade duration decreases in proportion to the increase in the ratio of the human order arrival rates with and without the presence of the machine. Trading volume goes up by the same rate. In other words, if the humans submit orders ten times faster when the machine is present, intertrade duration falls and trading volume increases by a factor of ten.

Second, we compute the optimal bid and ask prices for the machine that optimizes expected profits subject to an inventory constraint. The inventory constraint prevents the machine from carrying a significant open position to the next intra-human-trade period. The optimal bid and the ask for the machine are close to being symmetric around the mean value of the human orders, with the distance from the middle value being determined by the inventory constraint–the less concerned the machine is about the size of the remaining inventory, the closer its bid and the ask prices are to each other. The expected profit of an optimizing machine is increasing in both the variance and the arrival frequency of human orders.

Our two findings are interrelated; one the one hand, an optimizing machine is able to make positive expected profits by "sniping" out human orders somewhat away from the front of the book. On the other hand, execution of the "sniping" order submission strategy results in a transaction price density with bulges near the front and thinner outer tails. In fact, in a special case, the faster humans submit and vary their orders, the more profits the machine makes.

Our model has a number of limitations. First, it is not an equilibrium model of a limit order market like those of Parlour (1998), Foucault (1999), Biais, Martimor and Rochet (2000), Parlour and Seppi (2003), Foucault, Kadan and Kandel (2005), Goettler, Parlour, and Rajan (2005), Back and Baruch (2007), and Biais and Weill (2009), among others. These papers aim to derive the equilibrium price formation process. In order to cope with the large dimensionality of the state and action spaces of limit order markets, these studies use stylized models with many simplifying assumptions. In contrast to the equilibrium considerations, we study the formation of transaction prices given the distribution of orders, which we take to be exogenous over very short periods of time.

Second, our model is not a dynamic expected utility maximization model like those of Avellaneda and Stoikov (2008), Kuhn and Stroh (2009), and Rosu (2009). Those studies assume specific functional forms that govern the traders' preferences. We take the approach of modeling the order submission process over a very short period of time without specifying the traders' preferences or their optimization problems. Our model is essentially a stationary sequence of one-period models, and all that matters is what happens during the interval between two trades.

While these limitations make our results less satisfactory from the equilibrium analysis point of view, our approach is more pragmatic. Our results work for any possible (continuous) distribution of orders - equilibrium or otherwise. Thus, if a model comes up with the description of an equilibrium order submission process, we can plug it into our analysis and get the distributions of transaction prices and transaction times. Moreover, our results can be easily applied to the transaction-level data. We make no assumptions about the (unobservable) objectives of traders; we only make assumptions about their order submission processes.

Finally, to our knowledge, this is the first model to formally investigate the impact of high frequency trading on transaction prices, trading volume, and intertrade duration, as well as to characterize the profits of a high frequency trader in terms of the properties of low frequency traders.[3]

Our paper proceeds as follows. Section 2 studies the benchmark model without machines. Section 3 compares the benchmark model to the model in which an infinitely fast machine is present, and solves the optimization problem of the machine. Section 4 presents some empirical implcations of our results. Section 5 concludes.

2 Benchmark Model: Identical (Slow) Traders

2.1 A Single Intra-Trade Period

The model setup is as follows. There are infinitely many (slow) traders who submit limit orders into an electronic limit order book with the intent to buy or sell a single asset.

We make the following simplifying assumption:

Assumption 2.1 *Each order is for one unit of the traded asset only.*

We focus on a single intra-trade period, during which new buy and sell orders arrive into the limit order book, where $t = 0$ represents the beginning of the period.

Buy order prices are assumed to be represented by a sequence of random variables B_{t_n}, where t_n are Poisson arrival times of the buy orders, with intensity γ_B. Similarly, S_{s_m} represent sell order prices, and they arrive with intensity γ_S. We denote by μ_B, μ_S the maximum buy order price and the minimum sell order price, respectively, among those that are already resting in the book at the beginning of the interval, i.e, at $t = 0$.

The orders go out of the book either if they are executed or if they are canceled.[4] We denote by M_t^B the maximum of existing buy order prices and by m_t^S the minimum of existing sell order prices at time $t \geq 0$. Orders at time $t > 0$ consist of the resting orders as of $t = 0$ and the newly arrived orders.

We define the execution time of the next trade as

$$\tau := \inf\{t : M_t^B \geq m_t^S\}$$

At execution time τ, the transaction price P_τ is set at the maximum buy price M_τ^B if the trade was triggered by the sell order that came in at time τ; otherwise, the transaction price is set at the minimum sell price m_τ^S.

We introduce the following assumption, which presents a simple framework for studying the randomness of the buy and sell orders:

[3]Cont, Stoikov, and Talreja (2008) present a stochastic stochastic model for the continuous-time dynamics of a limit order book, but do not explicitly model high and low frequency traders.

[4]We will essentially assume away cancelations in what follows.

Assumption 2.2 *(i) Incoming buy orders B_{t_n} are iid with distribution F_B, conditionally on the information available by time $t = 0$. Similarly, incoming sell orders, S_{t_n}, are iid with distribution F_S. (ii) F_B and F_S have densities, and the densities are strictly positive for all x for which $0 < F_i(x) < 1$, $i = B, S$. (iii) Sell orders are independent from the buy orders. (iv) Within the intra-trade interval, the maximum buy order, M^B, and the minimum sell order, m^S, do not get canceled.*

We begin by computing the the distributions of bid and ask prices, and the distribution of the intra-trade time. We have the following result.

Proposition 2.1 *Under our standing assumptions, we have the following:*

(i) The distribution of the minimum sell order price among those that arrived by time t is given by

$$F_{m_t^S}(x) = 1 - \mathbf{1}_{\{x<\mu_S\}} e^{-t\gamma_S F_S(x)}$$

(ii) The distribution of the maximum buy order price among those that arrived by time t is given by

$$F_{M_t^B}(x) = \mathbf{1}_{\{x\geq\mu_B\}} e^{-t\gamma_B(1-F_B(x))}$$

(iii) Distribution of the time of trade is given by

$$P(\tau > t) = P(M_t^B < \mu_S)P(m_t^S = \mu_S) + \int_{\mu_B}^{\mu_S} P(M_t^B \leq x) dF_{m_t^S}(x)$$

$$= e^{-t\gamma_B(1-F_B(\mu_S))} e^{-t\gamma_S F_S(\mu_S)} + \int_{\mu_B}^{\mu_S} e^{-t\gamma_B(1-F_B(x))} t\gamma_S e^{-t\gamma_S F_S(x)} dF_S(x)$$

In particular, if the distributions of buy and sell order prices are the same, $F_B = F_S = F$, and γ_B is different from γ_S, then the distribution of the time of trade is given by

$$P(\tau > t) = e^{-t(\gamma_B[1-F(\mu_S)]+\gamma_S F(\mu_S))} + \frac{\gamma_S}{\gamma_B - \gamma_S} e^{-t\gamma_B} \left[e^{t(\gamma_B-\gamma_S)F(\mu_S)} - e^{t(\gamma_B-\gamma_S)F(\mu_B)}\right]$$

If, moreover, the new orders take values only inside the initial bid-ask spread, that is, $F(\mu_B) = 0$, $F(\mu_S) = 1$, then the distribution is that of the sum of two independent exponentials:

$$P(\tau > t) = \frac{\gamma_B}{\gamma_B - \gamma_S} e^{-t\gamma_S} - \frac{\gamma_S}{\gamma_B - \gamma_S} e^{-t\gamma_B}$$

If $F_B = F_S = F$ and $\gamma_B = \gamma_S = \gamma$, we get

$$P(\tau > t) = e^{-t\gamma}(1 + t\gamma[F(\mu_S) - F(\mu_B)])$$

with the mean

$$E[\tau] = \frac{1}{\gamma}[1 + F(\mu_S) - F(\mu_B)]$$

This proposition gives us the full description of the distributions of bid and ask prices, as well as that of the intra-trade time, as a functional of the distributions of buy and sell orders and their frequency. Thus, it also gives us information about the volume in a given interval of time. Interestingly, we see that in the symmetric case $F_B = F_S = F$ the time of trade distribution depends on F only through its values $F(\mu_B)$ and $F(\mu_S)$ evaluated at the initial bid and ask. Moreover, if also the new orders take values only inside the initial bid-ask spread, the expected time to trade is the sum of the expected buy and sell arrival times. Otherwise, the latter sum is the upper bound for the expected time to trade.

To illustrate this and subsequent results, we denote the range of buy order prices in the limit order book by $[A, B]$, and the range of sell order prices by $[C, D]$ where B and D can be infinite. In order to exclude uninteresting cases, we assume that

$$A \leq C \leq B \leq D \ , \mu_B \in [A, B] \ , \quad \mu_S \in [C, D]$$

Corollary 2.1 *Assume that F_B is uniform on $[A, B]$ and F_S is uniform on $[C, D]$, that $\gamma_B(D - C) \neq \gamma_S(B - A)$, and that the initial bid and ask can be ignored, that is, $\mu_B \leq C$, $\mu_S \geq B$. Then,*

$$P(\tau > t) = e^{-t\gamma_S \frac{B-C}{D-C}} + \frac{\gamma_S}{(D-C)[\frac{\gamma_B}{B-A} - \frac{\gamma_S}{D-C}]} \left[e^{-t\gamma_S \frac{B-C}{D-C}} - e^{-t\gamma_B \frac{B-C}{B-A}}\right]$$

with the mean

$$E[\tau] = \frac{\gamma_B(D - C) + \gamma_S(B - A)}{\gamma_S \gamma_B (B - C)}$$

The above corollary presents the expression for the expected time to trade in a special case when the distributions of buy and sell orders are assumed to be uniform. From the corrolary, the expected time to trade, $E[\tau]$, is large when B is close to C, so there is a small overlap between the possible values of buy and sell orders. Moreover, the expected time to trade is large if either the frequency for the arrival of buy orders or the frequency for the arrival of sell orders (or both) is low. In contrast, the expected time to trade can get shorter if the small overlap between the possible values for buy and sell orders can be made up for by an increase in the buy or sell frequency or if low order arrival frequency can be compensated by an increase in the buy-sell order overlap.

Next, we compute the distribution of transaction prices at a given time of trade, τ. We first introduce the probability that, conditional on an order arriving, it was a buy order:

$$p := \frac{\gamma_B}{\gamma_B + \gamma_S}$$

Denote by $A(\tau)$ the event that the order that just came in and triggered the transaction was a sell order, and by $A^c(\tau)$ the event that the transaction was triggered by an incoming buy order. The transaction price is defined as

$$P_\tau := M_\tau^B \mathbf{1}_{A(\tau)} + m_\tau^S \mathbf{1}_{A^c(\tau)}$$

Proposition 2.2 *Under our standing assumptions, we have the following:*

(i) The distribution of the maximum buy order price at the time of trade is given by, for $x \in [\mu_B, B]$,

$$P(M_\tau^B \leq x) = p(1-p)\times$$

$$\left[\int_{\mu_B}^{x \wedge \mu_S} \frac{F_B(x) - F_B(y)}{[(1-p)F_S(y) + p(1 - F_B(y))]^2} dF_S(y) + \int_{\mu_B}^{x \wedge \mu_S} \frac{F_S(y)}{[(1-p)F_S(y) + p(1 - F_B(y))]^2} dF_B(y)\right]$$

$$+[F_B(x) - F_B(\mu_S)]\frac{p}{p + (1-p)F_S(\mu_S) - pF_B(\mu_S)]}$$

$$+F_S(\mu_B)\frac{1-p}{p + (1-p)F_S(\mu_B) - pF_B(\mu_B)}$$

(ii) The distribution of the minimum sell price at the time of trade is given by, for $x \geq C$,

$$P(m_\tau^S \leq x) = p(1-p)\times$$

$$\left[\int_{\mu_B}^{x\wedge\mu_S} \frac{1 - F_B(y)}{[(1-p)F_S(y) + p(1 - F_B(y))]^2} dF_S(y) + \int_{\mu_B}^{x\wedge\mu_S} \frac{F_S(y\wedge x)}{[(1-p)F_S(y) + p(1-F_B(y))]^2} dF_B(y)\right]$$

$$+\mathbf{1}_{\{x>\mu_S\}}[1 - F_B(\mu_S)]\frac{p}{p + (1-p)F_S(\mu_S) - pF_B(\mu_S)}$$

$$+\mathbf{1}_{\{x>\mu_B\}}F_S(\mu_B\wedge x)\frac{1-p}{p + (1-p)F_S(\mu_B) - pF_B(\mu_B)}$$

(iii) The distribution of the transaction price is given by, for $x \in [\mu_B \vee C, D]$,

$$P(P_\tau \leq x) = p(1-p)\times\int_{\mu_B}^{x\wedge\mu_S}\frac{F_S(y)}{[(1-p)F_S(y) + p(1-F_B(y))]^2}dF_B(y)$$

$$+F_S(\mu_B)\frac{1-p}{p + (1-p)F_S(\mu_B) - pF_B(\mu_B)}$$

$$+p(1-p)\times\int_{\mu_B}^{x\wedge\mu_S}\frac{1 - F_B(y)}{[(1-p)F_S(y) + p(1-F_B(y))]^2}dF_S(y)$$

$$+\mathbf{1}_{\{x>\mu_S\}}[1 - F_B(\mu_S)]\frac{p}{p + (1-p)F_S(\mu_S) - pF_B(\mu_S)}$$

If $F_B = F_S = F$ *and* $p \neq 1/2$, *this becomes*

$$P(P_\tau \leq x) = \mathbf{1}_{\{x>\mu_B\}}\frac{p(1-p)}{1-2p}\left[\frac{1}{p + (1-2p)F(\mu_B)} - \frac{1}{p + F(x\wedge\mu_S)(1-2p)}\right]$$

$$+F(\mu_B)\frac{1-p}{p + (1-2p)F(\mu_B)}$$

$$+\mathbf{1}_{\{x>\mu_S\}}[1 - F(\mu_S)]\frac{p}{p + (1-2p)F(\mu_S)}$$

If, in addition to $F_B = F_S = F$, *we have* $p = 1/2$, *then we get*

$$P(P_\tau \leq x) = \mathbf{1}_{\{x>\mu_B\}}F(x\wedge\mu_S) + \mathbf{1}_{\{x>\mu_S\}}[1 - F(\mu_S)]$$

that is, $F_P = F$ *in the interval* (μ_B, μ_S).

The four terms in the price distribution given in (iii) are due to the following: the first two terms correspond to an incoming sell order being the new minimum and triggering the sale, where the second, non-integral term corresponds to the states of the world in which none of the buy orders that have arrived since the last trade is higher than the initial maximum buy order μ_B (so that the transaction price equals μ_B); the last two terms correspond to an incoming buy order being the new maximum and triggering the sale, where the very last, non-integral term corresponds to the states of the world in which none of the sell orders that have arrived since the last trade is lower than the initial minimum sell order μ_S (so that the transaction price equals μ_S).

In the case $F_S = F_B = F$, denoting by f the density of F, the density of the transaction price in the interval (μ_B, μ_S) is given by

$$f_{P_\tau}(x) = \frac{p(1-p)}{[p + F(x)(1-2p)]^2} f(x)$$

The factor multiplying $f(x)$ is increasing in $F(x)$ for $p > 0.5$. That is, if the buy orders are more likely, then the density f of order prices is distorted in favor of high transaction prices. The opposite is true if the sell orders are more likely.

Similarly, for a fixed and small value of $F(x)$, the factor multiplying $f(x)$ is decreasing in p – increasing p means less sell orders and more buy orders, so that the probability of the transaction price being small becomes lower. The opposte is true for a fixed and high value of $F(x)$.

Corollary 2.2 *Assume now that F_B is uniform on $[A, B]$ and F_S is uniform on $[C, D]$, that $\gamma_B(D - C) \neq \gamma_S(B - A)$, and that the initial bid and ask can be ignored, that is, $\mu_B \leq C$, $\mu_S \geq B$. Then, we have, for $x \in [C, B]$,*

$$P(P_\tau \leq x) = \frac{p(1-p)(B-C)}{p(D-C) - (1-p)(B-A)}$$

$$\times \left[\frac{1}{p\frac{B}{B-A} + (p-1)\frac{C}{D-C} + x\left(\frac{1-p}{D-C} - \frac{p}{B-A}\right)} - \frac{B-A}{p(B-C)} \right]$$

with the density

$$f_P(x) = p(1-p)\frac{(B-C)(B-A)(D-C)}{\{pB(D-C) + (p-1)C[B-A] + x[(1-p)(B-A) - p(D-C)]\}^2}$$

If, in addition, $D - C = B - A$, the expected value of the price is, in terms of the liquidity variable $z = \frac{p}{1-p} = \frac{\gamma_B}{\gamma_S}$,

$$E[P_\tau] = \frac{zB - C}{z - 1} - \frac{z}{(z-1)^2}(B - C)\log(z)$$

and the variance is

$$Var[P_\tau] = \frac{z(B-C)^2}{(z-1)^2}\left[1 - \frac{z}{(z-1)^2}(\log z)^2\right]$$

We see that with a loss of liquidity on the buy side (as $z \to 0$), the expected price tends to its lowest possible value C, and with the loss of liquidity on the sell side the expected price tends to its highest possible value B. In either case, the variance tends to zero. It can also be verified that the expected price is an increasing concave function of z, while the variance is a concave function of z with a maximum at $z = 1$.[5]

Remark 2.1 A digression on equilibrium order submission processes. We must caution the reader that although results obtained under the assumptions of symmetric distributions of buy and sell orders, $F_B = F_S = F$, are elegant and tractable, such symmetric order distributions may not arise in a full information equilibrium. In fact, we show in the appendix that, if the buyers are all identical, if they believe that the sellers follow the same distribution $F_S = F_B$ for their orders, and if they are risk-neutral, then the necessary condition for the buyers leads to the distribution of the form

$$F_B(x) = c_B(v_B - x)^{-2/3}$$

where c_B is a constant, and v_B is the value the buyers assign to holding one unit of the asset. [6] However, under the analogous assumptions on the sellers, the necessary condition for the sellers leads to the distribution of the form

$$F_S(x) = 1 - c_S(x - v_S)^{-2/3}$$

Thus, the assumption of the form $F_B = F_S = F$ is actually not tenable in such a full information equilibrium.

2.2 Multiple Periods

In order to extend our model to a multi-period setting, we need to specify how the orders change from one intra-trade period to another. We consider the case which will keep the setup as stationary as possible. We assume that, conditional on the last transaction price, $P(k)$, buy order $B_i(k+1)$ in the next intra-trade period is given by

$$B_i(k+1) = P(k) \times B_i$$

where B_i are iid with distribution F_B. In other words, the orders are equal to the previous price randomly distorted by a multiplicative random factor (which means the log-order is the previous log-price plus a random term). Set, without loss of generality, $P(0) = 1$ and denote $P(1) = P$. The (conditional) distribution of the buy orders in the $(k+1)$-st period is

$$F_{B,k+1}(x) = F_B(x/P(k))$$

and similarly for sell orders. Assume now that the book is emptied after the previous trade. It is then easily verified from Proposition 2.2 (iii) that in this model

$$F_{P(k+1)|P(k)}(x) = F_P(x/P(k)) \tag{2.1}$$

[5]The variable z is one measure of liquidity - the difference between the arrival frequencies of buy and sell orders. More broadly, liquidity reflects the ease with which an asset can be bought or sold without a significant effect on its price. Thus, liquidity has a number of other dimensions that are not being captured by z.

[6]For this to be a distribution function, x should take values less than a constant $v < v_B$.

where $F_{P(k+1)|P(k)}$ is the conditional distribution. In other words, we can write

$$P(k+1) = P \times P(k)$$

This means that the log-price is a random walk: it is obtained as a sum of iid random variables each with the distribution of $P = P(1)$. Thus, the distribution of the relative return is

$$P(\frac{P(k+1)}{P(k)} - 1) \leq x) = F_P(x+1)$$

So, under the assumptions of this section, in order to study the qualitative properties of the returns distribution, it suffices to study the price distribution.

Moreover, we also see that, denoting by τ_k the times of trade,

$$F_{\tau_{k+1}|P(k)}(t) = F_{\tau_k}(t)$$

and thus the intra-trade distribution is stationary.

3 A Model With A Machine Trader

The setup of the model is the same as in the benchmark model with the addition of one infinitely fast (from the point of view of other traders) high frequency trader. The high frequency trader–the machine–is assumed to keep issuing the same buy order b and sell order s, $b < s$, until a trade occurs. The orders b and s get immediately canceled if not executed right away. This mimics the so-called "sniping" strategy – a strategy designed to discover liquidity in the limit order book, or to "pick-off" orders already in the book.

We assume that the machine is so fast that it will always pick off a human sell order, S_i, before any other human trader whenever $b \geq S_i$, (machine buys for S_i), unless S_i is less than the existing maximum buy order μ_B in the book, in which case the transaction is executed at price μ_B. The assumption for the human buy orders, B_i, is similar.

Our objective is to compare a market with a machine to the one without it. Our comparison is non-strategic: in both setups, we maintain the same assumptions about the distributions and the arrival frequencies of the buy and sell orders. In a strategic setting, it is quite possible that the presence of a machine would affect order human submission processes and frequencies. However, we show that as in the benchmark case, the distribution of the transaction prices depends on the arrival rates only through the ratio $p = \gamma_B/(\gamma_B + \gamma_S)$. Thus, if the arrival rates change by the same factor because of the machine presence, p will not change, and there will be no effect on the price distribution. (On the other hand, there might be effects from the changes in the orders distributions.)

We now proceed to examine the distributions of execution times and transaction prices.

Proposition 3.1 *Assume $\mu_B < b < s < \mu_S$. The distribution of the time until next trade is given by*

$$P(\tau > t) = P(M_t^B < s)P(m_t^S \geq s) + \int_b^s P(M_t^B \leq x)dF_{m_t^S}(x)$$

$$= e^{-t\gamma_S F_S(s)} e^{-t\gamma_B(1-F_B(s))} + \int_b^s t\gamma_S e^{-t\gamma_S F_S(x)} e^{-t\gamma_B(1-F_B(x))} dF_S(x)$$

In particular, if $F_B = F_S = F$ and $\gamma_B = \gamma_S = \gamma$, we get

$$P(\tau > t) = e^{-t\gamma}[1 + t\gamma(F(s) - F(b))]$$

with the mean equal to

$$E[\tau] = \frac{1}{\gamma}[1 + F(s) - F(b)]$$

As we can see from the last expression, when $\mu_B < b < s < \mu_S$, $F_B = F_S = F$ and $\gamma_B = \gamma_S = \gamma$, denoting by γ_0 the arrival rates in the benchmark case with no machine, the ratio of the mean time between transactions with and without the machine is $\frac{\gamma}{\gamma_0}\frac{1+F(s)-F(b)}{1+F(\mu_S)-F(\mu_B)}$, which is less than $\frac{\gamma}{\gamma_0}$, but not less than half thereof. If the order arrival rates with the machine and without the machine are the same, i.e., $\gamma = \gamma_0$, then the presence of the machine speeds up the trades, but not more than by a factor of two, on average. By construction, the volume goes up, but not more than double.

If, however, the ratio $\frac{\gamma}{\gamma_0}$, is large, then the presence of the machine will speed up the trades (lower inter-trade duration) in proportion to this ratio. The volume will also increase by the same proportion.

We now present the result for the uniform distribution.

Corollary 3.1 *Assume $\mu_B < C < b < s < B < \mu_S$, that F_B is uniform on $[A, B]$ and F_S is uniform on $[C, D]$, and that $\gamma_B(D - C) \neq \gamma_S(B - A)$. Then, we have*

$$P(\tau > t) = e^{-t[\gamma_B\frac{B-s}{B-A} + \gamma_S\frac{s-C}{D-C}]} + \frac{\gamma_S}{D-C}\frac{e^{-t[\frac{\gamma_B B}{B-A} - \frac{\gamma_S C}{D-C}]}}{\frac{\gamma_B}{B-A} - \frac{\gamma_S}{D-C}}\left[e^{ts[\frac{\gamma_B}{B-A} - \frac{\gamma_S}{D-C}]} - e^{tb[\frac{\gamma_B}{B-A} - \frac{\gamma_S}{D-C}]}\right]$$

with the mean

$$E[\tau] = \frac{(B-A)(D-C)}{\gamma_B(D-C) - \gamma_S(B-A)}$$

$$\times\left[\frac{\gamma_B(D-C)}{\gamma_B(D-C)(B-s) + \gamma_S(B-A)(s-C)} - \frac{\gamma_S(B-A)}{\gamma_B(D-C)(B-b) + \gamma_S(B-A)(b-C)}\right]$$

In Figure 1 we plot the density of the intra-trade time with and without the machine for the uniform distribution of orders. In the presence of the machine, this density is more concentrated on low values.

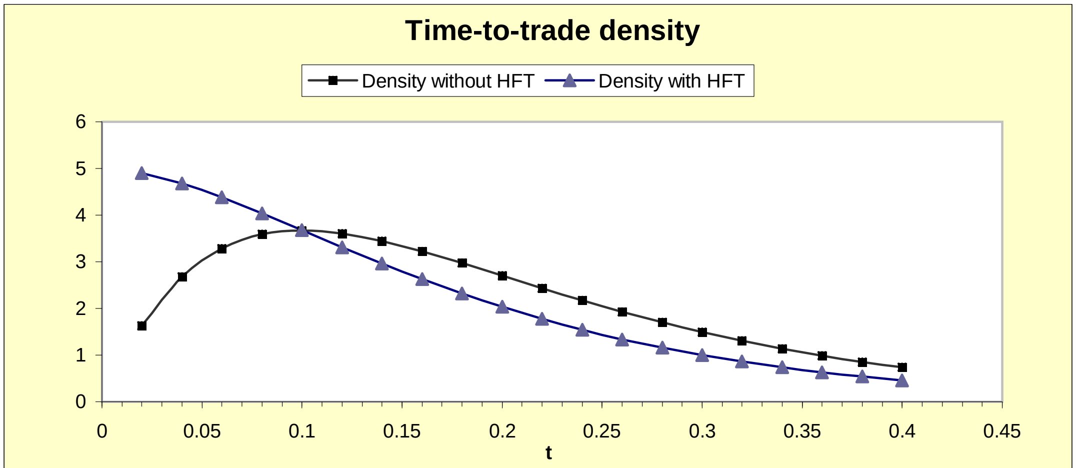


Figure 1: Density of intratrade time with and without the high-frequency trader.

Figure 2 shows the mean intra-trade times with and without the machine, as the supports of the buy and sell orders distributions have decreasing overlap. Average intertrade duration increases with less overlap, but the increase is steeper without machine.

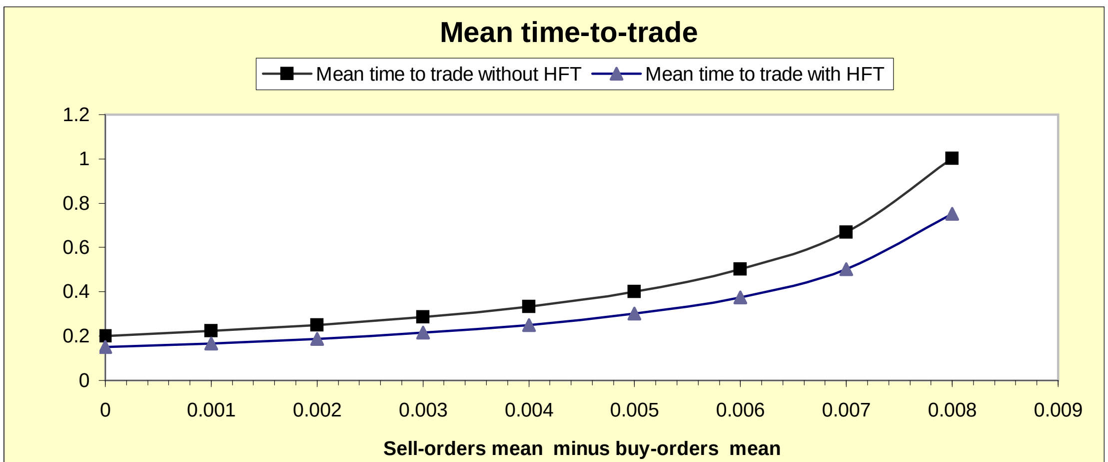


Figure 2: Mean intratrade time as the sell-buy gap increases.

The following is the main technical result of this section, while its economics consequences are given in the corollary below.

Proposition 3.2 *Assume $\mu_B < b < s < \mu_S$. The distribution of the transaction price P_τ is given by:*

$$P(P_\tau \leq x) = p(1-p)\times$$

$$\Big[\int_{\mu_B}^{x\wedge s} \frac{F_S(y)}{[(1-p)F_S(y\vee b)+p(1-F_B(y))]^2}dF_B(y)+\int_b^{x\wedge \mu_S}\frac{1-F_B(y)}{[(1-p)F_S(y)+p(1-F_B(y\wedge s))]^2}dF_S(y)\Big]$$

$$+\mathbf{1}_{\{x>\mu_B\}}F_S(\mu_B)\frac{1-p}{p+(1-p)F_S(b)-pF_B(\mu_B)}$$

$$+\mathbf{1}_{\{x>\mu_S\}}[1-F_B(\mu_S)]\frac{p}{p+(1-p)F_S(\mu_S)-pF_B(s)}$$

$$+p\times\mathbf{1}_{\{x>s\}}\int_s^{x\wedge\mu_S}\frac{1}{(1-p)F_S(y)+p(1-F_B(s))}dF_B(y)$$

$$+(1-p)\times\int_{\mu_B}^{x\wedge b}\frac{1}{(1-p)F_S(b)+p(1-F_B(y))}dF_S(y)$$

In particular, in case $F_B=F_S=F$, the price density on the interval (μ_B,μ_S) is given by

$$dF_P(x)=dF(x)\left[\mathbf{1}_{\{\mu_B<x<b\}}\frac{(1-p)[p+(1-p)F(b)]}{[(1-p)F(b)+p(1-F(x))]^2}+\mathbf{1}_{\{b\le x<s\}}\frac{p(1-p)}{[(1-2p)F(x)+p]^2}\right.$$
$$\left.+\mathbf{1}_{\{s\le x<\mu_S\}}\frac{p[1-p+p(1-F(s))]}{[(1-p)F(x)+p(1-F(s))]^2}\right]$$

and, when in addition $p=1/2$, on the interval (μ_B,μ_S) we have

$$dF_P(x)=dF(x)\left[\mathbf{1}_{\{\mu_B<x<b\}}\frac{1+F(b)}{[F(b)+1-F(x)]^2}+\mathbf{1}_{\{b\le x<s\}}+\mathbf{1}_{\{s\le x\le\mu_S\}}\frac{2-F(s)}{[F(x)+1-F(s)]^2}\right]$$

The following is the main economic result of this section, and it is obtained by direct examination of the price distribution given in the previous proposition, and the analogous result for the benchmark case of no machine. Here, we assume that the order distributions F_B,F_S and the probability of a buy order p are the same in the markets without and with the machine.

Corollary 3.2 *(i) Inside the interval $[b,s]$ the density of the transacted price remains the same as in the benchmark case. The far tails are more narrow, that is, the probabilities of the price being equal to μ_B and μ_S are lower, and, if μ_B is low enough, the density is lower for x greater than but close to μ_B, and analogously for x close to μ_S. The values of the density are higher at values less than but close to b and at values larger than but close to s.*

(ii) For a fixed price value $\mu_B<x<b_1<b_2$, its density $f_P(x)$ is higher if the machine uses lower bid b_1 than if it uses higher bid b_2, and analogously for $s_1<s_2<x<\mu_S$, the density is higher if the machine uses the higher ask s_2.

(iii) Assume now $F_B=F_S=F$ where F is symmetric, and $p=1/2$. If b and s are chosen symmetrically so that $F(b)=1-F(s)$, and the same is true for μ_B and μ_S, then the mean value of the transacted price is the same as the mean value of the incoming human orders, hence the same as the mean of the transacted price when there is no machine.

The intuition behind (i) is the following. The density remains the same on the interval (b,s) because the transaction price will take a value in that interval if and only if the transaction was between two human traders. Outside of this interval, but close to it, the density is higher relative to the benchmark case, as now the orders outside the interval $[b,s]$ get picked off by the machine. To compensate, the density has to go down in the far tails.

From (i) we see that the effect on the variance is complex – the thinning of the far tails would reduce the variance, but the fattening of the nearer parts of the tails has the opposite effect. Whether the variance goes up or down will depend on the actual values of b, s, μ_B,

μ_S, and on the distributions F_B, F_S. However, the higher even moments are likely to go down, because of the thinning of the far tails. Furthermore, it can be verified that, if the ratio $f_S(x)/f_B(x)$ of the sell vs. buy order densities is bounded from above and away from zero, then also bounded is the ratio of the density of the transaction price with machine vs. that density without machine. Also, what we have just discussed is the variance of a single transaction price. Let us recall that the time between transactions goes down in the presence of the machine (at most by a factor of two). Thus, even if the variance of the single transaction price goes up, the variance of the average transaction price per unit time may go down.

The first part of item (ii) holds because there is higher density for values between b_1 and b_2 if the machine uses b_2, as it picks off those values, too. Thus, to compensate for this, the density has to go down for values of x below b_1 (the machine picks off fewer of those). Similarly on the ask side.

Item (iii) gives conditions under which the mean price does not change. Perhaps more interestingly, if these conditions are not satisfied, the mean price is likely to change, in general. Thus, the presence of the sniping machine is likely to change the average transaction price, even in the absence of new information, if the distributions of the sell orders and buy orders are not symmetric, or if the machine's bid and ask are not symmetric with respect to the orders distribution.

In the case of the uniform distribution we get

Corollary 3.3 *Assume $\mu_B < C < b < s < B < \mu_S$, that F_B is uniform on $[A, B]$ and F_S is uniform on $[C, D]$, and that $\gamma_B(D - C) \neq \gamma_S(B - A)$. Then, the density of the price for $x \in [C, B]$ is given by*

$$f_P(x) = \mathbf{1}_{\{x<b\}} \frac{p(1-p)(B-C)(B-A)(D-C) + (1-p)^2(B-A)^2(b-C)}{[(1-p)(b-C)(B-A) + p(B-x)(D-C)]^2}$$

$$+\mathbf{1}_{\{b<x<s\}} \frac{p(1-p)(B-C)(B-A)(D-C)}{[(1-p)(x-C)(B-A) + p(B-x)(D-C)]^2}$$

$$+\mathbf{1}_{\{x>s\}} \frac{p(1-p)(B-C)(B-A)(D-C) + p^2(D-C)^2(B-s)}{[(1-p)(x-C)(B-A) + p(B-s)(D-C)]^2}$$

Figure 3 illustrates the conclusions of Corollary 3.2, showing the thinning of the far tails of the density, the fattening for the values moderately away from the middle of the distribution, and no change in the middle.

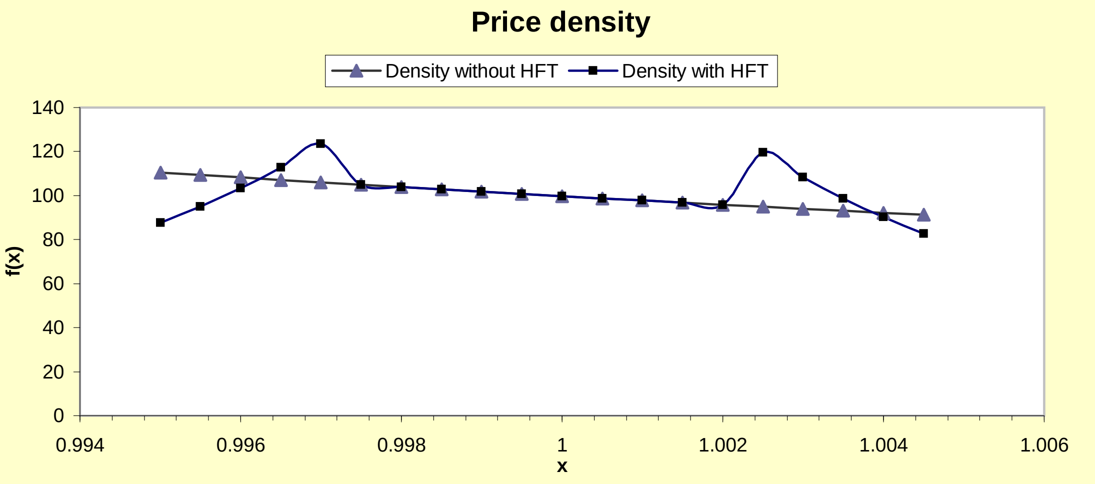


Figure 3: Price density when orders are uniform.

Figures 4 and 5 show the means and the variances of the price with and without machine presence, as the supports of the uniform distributions of orders have less and less overlap.[7]

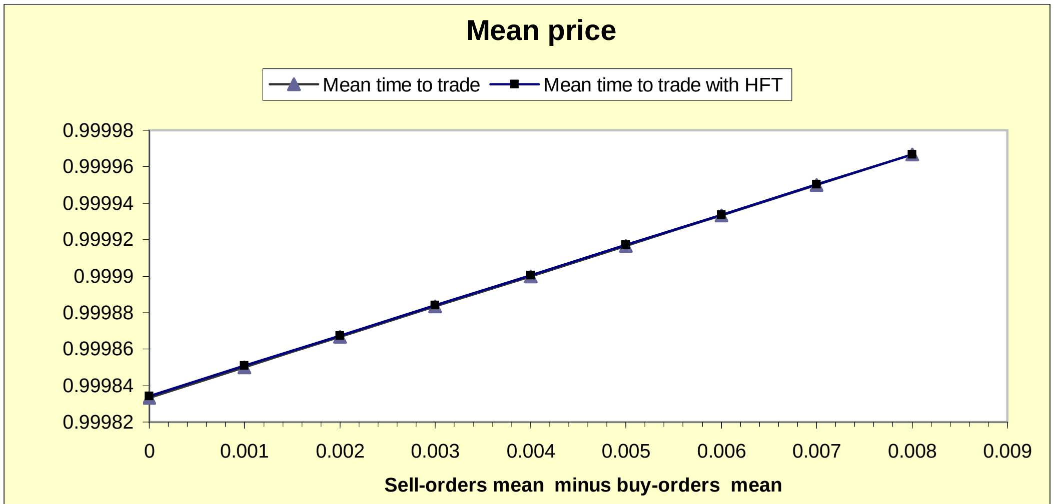


Figure 4: Mean price as the sell-buy gap increases.

[7]We decrease the overlap by moving to the right the support interval for the sell orders and keeping the same the distribution of the buy orders.

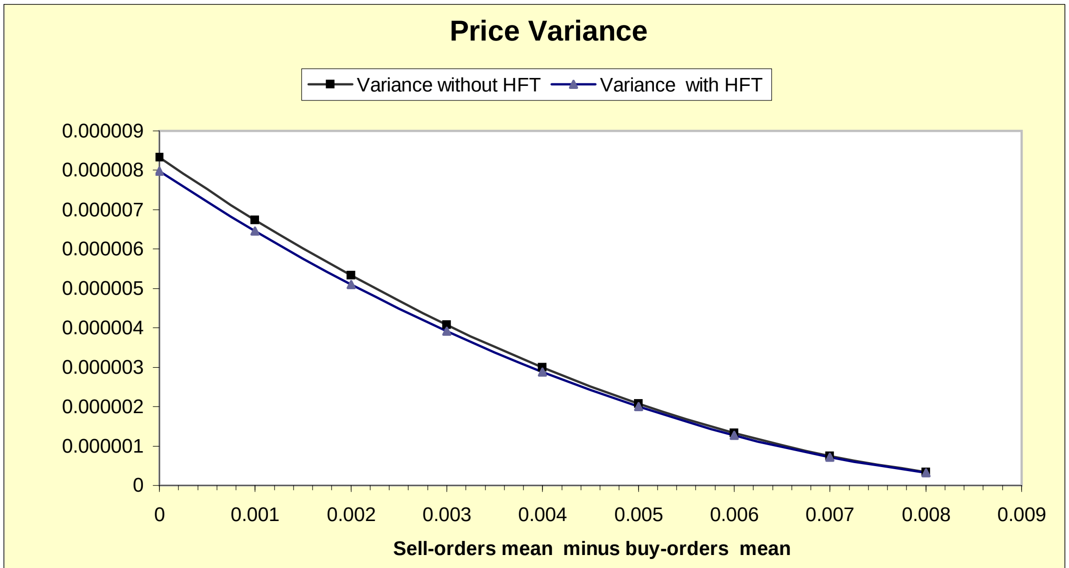


Figure 5: Price variance as the sell-buy gap increases.

The average values are almost identical in the two cases, while the variance with machine is somewhat lower than without it, but the difference vanishes as the supports of the buy and sell orders diverge.

3.1 Machine Optimization

Up to now, we assumed that the machine submits very fast buy and sell orders and cancels them if they are not executed. Under this assumption, machine does not learn from transaction prices or the execution of orders.

Let us now assume that the machine will be issuing the same orders b and s until a random time τ, which is less or equal to the first time a human order "steals" from the machine a human sell order $S_i < b$ or a human buy order $B_i > s$. The machine interprets the time τ as the first time some new information arrives in the market. For simplicity, we assume that over very short intervals of time that we focus on, the machine models τ as a random time independent of everything else, having exponential distribution with intensity λ.[8] Also for simplicity, we set $\mu_B = 0$, $\mu_S = \infty$, that is, the book is initially empty.

Denote by N_b (N_s) the number of buys (sells) of the machine during the random period $[0, \tau]$. Also denote

$$p_S = P(b \geq S_i), \quad p_B = P(s \leq B_i)$$

$$r_S = E[S_i \mathbf{1}_{b \geq S_i}] \ , \quad r_B = E[B_i \mathbf{1}_{s \leq B_i}]$$

Note that N_b, N_s are conditionally binomial with probability p_S, p_B, and the number of trials being Poisson with intensity γ_S, γ_B.

[8]If we require that τ is less or equal to the first "stealing time", then it is not really independent of everything, but we assume that the machine uses independence as an approximating assumption.

Lemma 3.1 *We have*

$$E[N_b] = \frac{1}{\lambda} p_S \gamma_S$$

$$E[N_s] = \frac{1}{\lambda} p_B \gamma_B$$

and the expected profit from buying and selling, ignoring the value of inventory, is

$$E[P] = \frac{1}{\lambda} r_B \gamma_B - \frac{1}{\lambda} r_S \gamma_S$$

Moreover, we have

$$E[N_b^2] = \frac{1}{\lambda} \gamma_S p_S + \frac{2}{\lambda^2} \gamma_S^2 p_S^2$$

$$E[N_s^2] = \frac{1}{\lambda} \gamma_B p_B + \frac{2}{\lambda^2} \gamma_B^2 p_B^2$$

$$E[N_b N_s] = \frac{2}{\lambda^2} \gamma_S p_S \gamma_B p_B$$

We suppose that the machine trader maximizes expected profit/loss during the interval, but penalized by the size of the inventory, and adjusted by the value of the remaining inventory. More precisely, the machine maximizes

$$E[G] := E[P] - \rho E[(N_s - N_b)^2] + v E[N_b - N_s]$$

where ρ is a penalization parameter, or a Lagrange multiplier for the inventory constraint, and v can be thought of as proportional to the estimated future value of the asset.

This problem is hard in general, and we only consider the case when the human orders are uniformly distributed.

3.1.1 Uniformly distributed orders

Let us assume uniform distributions

$$F_B(x) = \frac{x - A}{B - A}\ , \quad F_S(x) = \frac{x - C}{D - C}$$

that is, B_i, S_i are respectively uniform on $[A, B]$, $[C, D]$.

Lemma 3.2 *If F_B is uniform on $[A, B]$ and F_S is uniform on $[C, D]$, then we have*

$$p_S = F_S(b) = \frac{b - C}{D - C}$$

$$p_B = 1 - F_B(s) = \frac{B - s}{B - A}$$

$$r_S = F_S(b)[C + F_S(b)(D - C)/2] = \frac{b^2 - C^2}{2(D - C)}$$

$$r_B = (1 - F_B(s))[B - (1 - F_B(s))(B - A)/2] = \frac{B^2 - s^2}{2(B - A)}$$

Proposition 3.3 *For maximizing $E[G]$, the interior first order condition with respect to s is*

$$s[1+\rho\frac{4}{\lambda}\gamma_B\frac{1}{B-A}] = \rho + 4\rho\frac{\gamma_B}{\lambda}\frac{B}{B-A} - 4\rho\frac{\gamma_S}{\lambda}\frac{b-C}{(D-C)} + v$$

The interior first order condition with respect to b is

$$b[1+\rho\frac{4}{\lambda}\gamma_S\frac{1}{D-C}] = -\rho + 4\rho\frac{\gamma_S}{\lambda}\frac{C}{D-C} + 4\rho\frac{\gamma_B}{\lambda}\frac{B-s}{(B-A)} + v$$

In particular, if

$$A = C, B = D, \gamma_B = \gamma_S = \gamma$$

then, the interior solutions are

$$s = \frac{4\rho\gamma(A+B) + v\lambda(B-A)}{\lambda(B-A) + 8\rho\gamma} + \rho$$

$$b = \frac{4\rho\gamma(A+B) + v\lambda(B-A)}{\lambda(B-A) + 8\rho\gamma} - \rho$$

Introducing the mean and the variance of the human orders,

$$\mu = (A+B)/2\ , \quad \sigma^2 = (B-A)^2/12$$

we get

$$s = \frac{8\rho\gamma\mu + v\sqrt{12}\lambda\sigma}{\sqrt{12}\lambda\sigma + 8\rho\gamma} + \rho$$

$$b = \frac{8\rho\gamma\mu + v\sqrt{12}\lambda\sigma}{\sqrt{12}\lambda\sigma + 8\rho\gamma} - \rho$$

From this proposition we find that (assuming interior solutions) the machine places orders centered around the mid-price $\frac{8\rho\gamma\mu+v\sqrt{12}\lambda\sigma}{\sqrt{12}\lambda\sigma+8\rho\gamma}$ adjusted for the inventory penalty ρ. This mid-price is less than the mean value of the incoming orders μ when the weight v given to the expected future asset value is small, and is otherwise larger than μ. In addition, when $\rho = 0$, then the optimal orders are simply $b = s = v$. Furthermore, when the trading interval until the time of new information gets longer (λ closer to zero), then the machine orders get closer to $\mu \pm \rho$. The same happens when the frequency γ of human orders gets large, or when the variance σ^2 of human orders gets small. When γ gets small, the orders get close to $v \pm \rho$.

3.1.2 Orders symmetric around the mean

We again assume $A = C$, $B = D$, $\gamma_b = \gamma_S$. Everything simplifies if we only allow the orders of the form

$$b = \mu - x\ , \quad s = \mu + x$$

As discussed above, this is close to optimal if the product $\lambda\sigma$ is small relative to the product $\rho\gamma$. Moreover, as stated below, with this choice the expected inventory size is zero, $E[N_b -$

$N_s] = 0$. Thus, the machine does not have to worry, in expected value sense, about the future value of the asset.

If we optimize over x, it is easily seen that it is optimal to take

$$x = \rho.$$

Interpreting now ρ as a Lagrange multiplier, assume now we impose a constraint on the inventory size as follows:

$$E[(N_s - N_b)^2] \leq K \tag{3.1}$$

The following result is easy to verify.

Proposition 3.4 *Under our assumptions, we have*

$$0 = p_S - p_B = E[N_b - N_s]$$

and thus

$$E[(N_s - N_b)^2] = \frac{\gamma}{\lambda}(p_S + p_B)$$

Moreover, the equality in (3.1) *will be attained for* ρ *given by*

$$\rho = \frac{1}{2}[1 - \frac{\lambda}{\gamma}K](B - A)$$

In particular,

$$b = \mu - \rho \geq A$$

Furthermore, the expected profit can be computed as

$$E[P] = \frac{\sqrt{3}\sigma}{2}K(2 - \frac{\lambda}{\gamma}K).$$

The highest inventory is attained for $\rho = 0$ which gives $K = \gamma/\lambda$. Thus, it suffices to consider the values $K < \gamma/\lambda$. For K, it may be reasonable to take

$$K = \frac{N^2}{\lambda^2}$$

where N is a given constant that represents the maximal allowed inventory size per unit time.

The proposition above states that the machine's profit is a linear increasing function of the human orders' volatility σ (in the domain $K < \gamma/\lambda$). In addition, the machine's profit is increasing in the frequency of human orders γ.

The machine's profit is bounded by $\frac{\sqrt{3}\sigma}{2}\frac{\gamma}{\lambda}$. Thus, if there were increasingly many machines, as the total profit would have to be shared, the profit for each one would be decreasing.

If we constrain the absolute size of inventory rather than its size per unit time, that is, if K is kept fixed in a way that it does not depend on λ, the machine's profit is increasing in the mean length of the trading interval $1/\lambda$. This is because more trades are likely to be executed. This is also the case if we limit the inventory size per unit time, that is, K

is proportional to $1/\lambda^2$, but only in the domain consisting of $1/\lambda$ small enough. For large enough $1/\lambda$, if the inventory per unit time is limited, then the machine's profit is decreasing in $1/\lambda$.

From the expression for ρ we conclude that the machine provides less liquidity, in the sense that ρ is larger, in the following cases: 1) the market is more volatile so that the volatility of the orders is larger (that is, $B - A$ is larger); 2) the humans are trying harder to change their positions, that is, frequency γ is higher; 3) the value of K does not depend on λ and there is less new information coming in, that is, λ is lower; 4) the value of K is proportional to $1/\lambda^a$ for $a > 2$ and there is more new information coming in, that is λ is higher. Note that item 1) implies, supposing that in the time of crisis the volatility and the frequency of orders are higher, and supposing λ does not change much, that the machine will provide less liquidity (wider bid-ask spread) when there is crisis.

Finally, we remark that if humans had perfect knowledge about the machine's strategy, then the humans would submit only orders with values inside the interval $[b, s]$. If they did this by choosing values from a continuous distribution on $[b, s]$, the machine would not be able to make any trades, and would have zero profit. On the other hand, with this knowledge it might be optimal for humans to submit orders with values b or s with positive probability, which would place us outside of the assumptions of our model. However, because b and s can change from one intra-trade interval to another, it is unlikely that humans would be able to know their exact values.

4 Empirical Implications

Our results have a number of empirical implications. First, the distribution of transaction prices (and returns) in markets with high frequency traders can be represented as a "mixture" of the distributions of human-human and machine-human transaction prices (plus machine-machine prices, if there is more than one machine). With the knowledge of counterparties for each transaction, one can reconstruct the mixture. In addition, if machine strategies can, indeed, be closely approximated by a deterministic process (e.g., bracketing the last human transaction price), then the component of the price distribution attributed to the machines should be forecastable. As the proportion of transactions with the machines grows, forecastability of transaction prices should improve.[9]

Second, trading volume and intertrade duration, as well as measures of market liquidity based on them, should increase in direct proportion to how much humans change the speed of their orders when the machine is present. To the extent that it is known how many order per unit time have been submitted (modified or canceled) by machines and humans, this implication can be verified in the data.

Third, profits of a high frequency trader should increase in both the variance and the arrival frequency of human orders. Again, to the extent that both the arrival frequency and the variance of human orders can be estimated, they can be empirically compared to the profits and losses of a high frequency trader, as well as these traders as a group.

[9] Even if that component of the prices is forecastable, this does not mean that one can trade on it. It would take a machine that's faster than the fastest current machine to take advantage of this empirical regularity.

5 Conclusion

We model an electronic limit order market populated by low frequency traders and then add a high frequency trader. We postulate that low frequency traders (humans) follow certain order submission strategies and then derive the distributions of transaction prices with and without a high frequency trader (the machine).

We find that the presence of a machine is likely to change the average transaction price, even in the absence of new information. We also find that in a market with a high frequency trader, the distribution of transaction prices has more mass around the center and thinner far tails. With a machine, mean intertrade duration decreases in proportion to the increase in the ratio of the human order arrival rates with and without the presence of the machine; trading volume goes up by the same rate.

We also find that a machine that optimizes expected profits subject to an inventory constraint submits orders that are essentially symmetric around the mean value of the human orders. The distance between the machine's bid and ask prices increases with its concern about the size of the remaining inventory. The expected profit of an optimizing machine increases in both the variance and the arrival frequency of human orders.

Our model has two serious limitations. First, we do not solve for mutually best responses of all parties; in other words, the order submission strategy that we postulate for the humans may or may not be supported as an equilibrium strategy under general conditions. Second, our model is static; we focus on a stationary sequence of one-period models–intervals between two human trades.

Having said that, to our knowledge, this is the first model to formally investigate the impact of high frequency trading on transaction prices, trading volume, and intertrade duration, as well as to characterize the profits of a high frequency trader as a function of the properties of low frequency traders.

References

[1] Avellaneda, M. and Stoikov,S. (2008), “High-frequency trading in a limit order book”. *Quantitative Finance*, 8, 217-224.

[2] Back, K. and Baruch, S. (2007), “Working Orders in Limit-Order Markets and Floor Exchanges”. forthcoming in the *Journal of Finance*, .

[3] Biais, B., Martimort, D. and Rochet, J.-C. (2000), “Competing mechanisms in a common value environment”. *Econometrica*, 68, 799837.

[4] Biais, B. and Weill, P.-O. (2009), “Liquidity shocks and order book dynamics”. Working paper.

[5] Cont, R., Stoikov, S. and Talreja, R. (2009), “A stochastic model for order book dynamics”. Working paper.

[6] Foucault, T. (1999), “Order flow composition and trading costs in a dynamic limit order market”. *Journal of Financial Markets*, 2, 99134.

[7] Foucault, T., Kadan, O. and Kandel, E. (2005), “Limit order book as a market for liquidity”. *Review of Financial Studies*, 18, 1171-1217.

[8] Kuhn, C. and Stroh, M. (2009), “Optimal portfolios of a small investor in a limit order market a shadow price approach”. Working paper.

[9] Rosu, I. (2009), “A dynamic model of the limit order book”. Forthcoming in *The Review of Financial Studies”*.

[10] Parlour, C. (1998), “ Price dynamics in limit order markets”. *Review of Financial Studies*, 11, 789-816.

[11] Parlour, A.S. and Seppi, D.J. (2003), “Liquidity-based competition for order flow”. *Review of Financial Studies*, 16, 301-343.

[12] Goettler, R., Parlour, C., and U. Rajan (2005), “Equilibrium in a dynamic limit order market”, *Journal of Finance*, 60, 2149-2192.

Appendix

Proofs for Section 2

Proof of Proposition 2.1: Conditioning on the number of sell orders, we get

$$F_{m_t^S}(x) = \sum_{k=o}^{\infty}[1 - \mathbf{1}_{\{x<\mu_S\}}(1 - F_S(x))^k]\frac{(t\gamma_S)^k}{k!}e^{-t\gamma_S}$$

which proves the result. Similarly for $F_{M_t^B}$.

Next,

$$P(\tau > t) = P(M_t^B < m_t^S)$$

which gives the desired expression.

If $F_S = F_B = F$, then the integral can be easily computed explicitly to get the result.

■

Proof of Proposition 2.2: Denote by K_τ^B, K_τ^S the number of newly arrived buy and sell orders in the book at the time of trade, by $M^B(r)$, $m^S(q)$ the maximum of r buy orders and μ_B, and the minimum of q sell orders and μ_S, and by $B(r)$, $S(q)$ the $r-$th incoming buy order and the $q-$th incoming sell order. Let us also denote $B(r,q)$ the event that, given that there are r buy orders (plus μ_B) in the buy side of the book, and q sell orders (plus μ_S) in the sell side of the book, the last order was a buy order. Similarly for $S(r,q)$, except the sell order was the last. Then, we have

$$P(B(r,q)) = \binom{r-1+q}{q} p^r(1-p)^q$$

$$P(S(r,q)) = \binom{r-1+q}{r} p^r(1-p)^q$$

Notice that we have

$$\sum_{q=0}^{\infty}\binom{r-1+q}{q}(1-p)^q = p^{-r}$$

$$\sum_{r=0}^{\infty}\binom{r-1+q}{r}p^r = (1-p)^{-q}$$

Also note that we have

$$P(m^S(q) \le x) = 1 - \mathbf{1}_{\{x<\mu_S\}}(1 - F_S(x))^q$$

so that

$$dF_{m^S(q)}(x) = \mathbf{1}_{\{x<\mu_S\}}q(1 - F_S(x))^{q-1}dF_S(x)$$

and similarly

$$P(M^B(r) \le x) = \mathbf{1}_{\{x\ge\mu_B\}}(F_B(x))^r$$

$$dF_{M^B(r)}(x) = \mathbf{1}_{\{x\ge\mu_B\}}r(F_B(x))^{r-1}dF_B(x)$$

Using the above, we have

$$P(M_\tau^B \le x) = \sum_{r,q} P(M_\tau^B \le x, K_\tau^B = r, K_\tau^B = q)$$

$$= \sum_{r\ge 1,q\ge 0} P(M_\tau^B \le x, K_\tau^B = r, K_\tau^B = q, B(r,q)) + \sum_{r\ge 0,q\ge 1} P(M_\tau^B \le x, K_\tau^B = r, K_\tau^B = q, S(r,q))$$

$$= \sum_{r\ge 1,q\ge 0} P(B(r,q))P(M^B(r) \le x, M^B(r-1) \le m^S(q) \le B(r))$$

$$+ \sum_{r\ge 0,q\ge 1} P(S(r,q))P(S(q) < M^B(r) \le \min\{x, m^S(q-1)\})$$

Conditioning on $m^S(q)$ in the first term and on $M^B(r)$ in the second term, we get

$$\begin{aligned} P(M_\tau^B \le x) \;=\; & \mathbf{1}_{\{x>\mu_B\}} \sum_{r\ge 1,q\ge 0} \int_{\mu_B}^{x\wedge\mu_S} P(B(r,q))F_B^{r-1}(y)[F_B(x)-F_B(y)]q(1-F_S(y))^{q-1}dF_S(y) \\ & +\mathbf{1}_{\{x>\mu_S\}} \sum_{r\ge 1,q\ge 0} P(B(r,q))P(m^S(q)=\mu_S)F_B^{r-1}(\mu_S)[F_B(x)-F_B(\mu_S)] \\ & + \sum_{r\ge 0,q\ge 1} \int_{\mu_B}^{x\wedge\mu_S} P(S(r,q))F_S(y)[1-F_S(y)]^{q-1} rF_B^{r-1}(y)dF_B(y) \\ & +\mathbf{1}_{\{x>\mu_B\}} \sum_{r\ge 0,q\ge 1} P(S(r,q))P(M^B(r)=\mu_B)[1-F_S(\mu_B)]^{q-1}F_S(\mu_B) \end{aligned}$$

Inside the first integral we have a sum of the form

$$\sum_{r,q} \binom{r-1+q}{q} q x^{q-1} y^{r-1}$$

This is a derivative with respect to x of the sum

$$\sum_{r,q} \binom{r-1+q}{q} x^q y^{r-1} = \sum_r \frac{y^{r-1}}{(1-x)^r}$$

Thus, taking the derivative, we get the sum

$$\frac{1}{(1-x)^2}\sum_r r\left(\frac{y}{1-x}\right)^{r-1} = \frac{1}{(1-x)^2}\frac{1}{(1-\frac{y}{1-x})^2}$$

Setting

$$x = (1-p)(1-F_S(y))\ ,\ \ y = pF_B(y)$$

we get the result for the first integral in the distribution of M_B. The second integral is obtained in a similar manner, and the second and fourth non-integral terms are obtained by direct summation, taking into account that

$$P(M^B(r)=\mu_B) = F_B^r(\mu_B)\ ,\quad P(m^S(q)=\mu_S) = [1-F_S(\mu_S)]^q$$

Similarly, we have

$$P(m^S_\tau \le x) = \sum_{r,q} P(m^S_\tau \le x, K^B_\tau = r, K^B_\tau = q)$$

$$= \sum_{r\ge 1, q\ge 0} P(B(r,q), M^B(r-1) \le m^S(q) \le x \wedge B(r))$$

$$+\mathbf{1}_{\{x>\mu_S\}} \sum_{r\ge 1, q\ge 0} P(B(r,q))P(m^S(q) = \mu_S)F_B^{r-1}(\mu_S)[1-F_B(\mu_S)]$$

$$+ \sum_{r\ge 0, q\ge 1} P(S(q) \le x \wedge M^B(r), M^B(r) \le m^s(q-1), S(r,q))$$

$$+ \sum_{r\ge 0, q\ge 1} P(S(r,q))P(M^B(r) = \mu_B)[1-F_S(\mu_B)]^{q-1}F_S(x \wedge \mu_B)]$$

Similarly as above, conditioning on $m^S(q)$ in the first two terms and on $M^B(r)$ in the last two terms, and by summation, we get the result.

The distribution of the transaction prices is now easily determined as above, from its definition.

■

Derivations related to Remark 2.1

Let x be a submitted buy order. Denote

$$p_1^B(x) = P(x \text{ executed at arrival})$$

$$p_2^B(x) = P(x \text{ executed after arrival})$$

and similarly $p_i^S(x)$ if x is a sell order.

Let us suppose that a buy trader has the value v for the asset, at the time of the next trade. Given that the trader submits a buy order x, and his utility function is U, denoting by τ_x the time of his arrival and recalling that m^S_t denotes the minimum sell order in the book at time t, his expected profit is

$$p_1 E[U(v - m^S_{\tau_x}) \mid x \text{ executed at arrival}] + p_2 U(v-x)$$

Let us now change the variables to

$$u := F_B(x)$$

and denote by $\pi(u)$ the corresponding expected profit. Then, u is uniformly distributed, and a necessary condition for a symmetric equilibrium is that $\pi'(u) = 0$.

Proposition 5.5 *We have*

$$p_1^B(x) = \int_0^x dF_S(y) \frac{\gamma_S\gamma_B}{[2\gamma_B + \gamma_S F_S(y) - \gamma_B F_B(y)]^2}$$

$$p_2^B(x) = \frac{\gamma_B\gamma_S F_S(x)}{\{2\gamma_B + \gamma_S F_S(x) - \gamma_B F_B(x)\}\{\gamma_B[1-F_B(x)] + \gamma_S F_S(x)\}}$$

In case $F_S = F_B$, $\gamma_S \neq \gamma_B$, and in terms of variable u, we can write (with a slight abuse of notation p_1),

$$p_1^B(u) = \frac{\gamma_S\gamma_B}{\gamma_S - \gamma_B}(\frac{1}{2\gamma_B} - \frac{1}{2\gamma_B + (\gamma_S - \gamma_B)u})$$

We also have

$$p_1^S(x) = \int_x^\infty dF_B(y)\frac{\gamma_S\gamma_B}{[\gamma_B + \gamma_S + \gamma_S F_S(y) - \gamma_B F_B(y)]^2}$$

$$p_2^S(x) = \frac{\gamma_B\gamma_S[1 - F_B(x)]}{\{\gamma_B + \gamma_S + \gamma_S F_S(x) - \gamma_B F_B(x)\}\{\gamma_B[1 - F_B(x)] + \gamma_S F_S(x)\}}$$

Moreover, the density of the minimum of sell orders at time of arrival of buy order x, denoted f_x^m, lives on $(0, x)$, and is given by

$$f_x^m(z) = \mathbf{1}_{\{x>z\}}\frac{(p_1^B)'(z)}{p_1^B(x)}$$

Similarly, the density of the maximum of sell orders at time of arrival of sell order x, denoted f_x^M, lives on (x, ∞), and is given by

$$f_x^M(z) = -\mathbf{1}_{\{z>x\}}\frac{(p_1^S)'(z)}{p_1^S(x)}$$

Let us assume linear utility for the buyer. Denote by A_1 the event that buy order x is executed at arrival. Then, the trader's utility is

$$p_1^B(x)E[v - m_{\tau_x}^S|A_1] + (v_B - x)p_2^B(x) = p_1^B(x)v_B - \int_0^x z(p_1^B)'(z)dz + (v_B - x)p_2^B(x)$$

If $F_B = F_S$, we get

$$p_1^B(x) = F(x)/4\ , \quad p_2^B(x) = F(x)/2$$

In terms of the variable $F(x) = u$, denoting by β the inverse of F, this then becomes

$$\frac{1}{4}[v_B u - \int_0^{\beta(u)} zF'(z)dz] + \frac{1}{2}u[v_B - \beta(u)]$$

Taking derivative, setting it equal to zero, solving the obtained ODE for $\beta(u)$ and inverting, we get

$$F(x) = c(v_B - x)^{-2/3}$$

A similar computation for the seller gives

$$F(x) = 1 - c(x - v_S)^{-2/3}$$

Thus, the assumption of $F_S = F_B$ is not tenable in this equilibrium.

Proof of Proposition 5.5: The execution probability can be decomposed over the event that the order was executed as as soon as it arrived, and over the event that it was executed later, but before any other order was executed. The probability over the first event, by conditioning over the time t of the arrival of x, over the number r of buy orders by time

t and the number q of sell orders by time t, and over the minimum y of those q sell orders, is given by

$$
\begin{aligned}
p_1 &:= P(x \text{ executed at arrival}) \\
&= \sum_{r\geq 0, q\geq 1} \int_0^\infty \gamma_B e^{-\gamma_B t} dt \int_0^x q(1-F_S(y))^{q-1} dF_S(y) e^{-\gamma_B t} \frac{(\gamma_B t)^r}{r!} e^{-\gamma_S t} \frac{(\gamma_S t)^q}{q!} F_B^r(y) \\
&= \sum_{q\geq 1} \int_0^\infty \gamma_B e^{-\gamma_B t} dt \int_0^x q(1-F_S(y))^{q-1} dF_S(y) e^{-\gamma_B t[1-F_B(y)]} e^{-\gamma_S t} \frac{(\gamma_S t)^q}{q!} \\
&= \int_0^\infty \gamma_S \gamma_B t e^{-\gamma_B t} dt \int_0^x dF_S(y) e^{-\gamma_B t[1-F_B(y)]} e^{-\gamma_S t F_S(y)} \\
&= \int_0^x dF_S(y) \frac{\gamma_S \gamma_B}{[\gamma_B(2-F_B(y)) + \gamma_S F_S(y)]^2}
\end{aligned}
$$

Denote now, similarly as before, by $S(r,q)$ the event that at the execution of order x, there were r buy orders and q sell orders that arrived after x, which necessarily implies that $q \geq 1$, and that the last order that arrived was a sell order. The probability of execution of x at some time after it arrived can be obtained by conditioning over the time t of the arrival of x, over the number k of buy orders by time t and the number l of sell orders by time t, and on the number r of buy orders and the number q of sell orders that arrived between time t and the time of execution. This probability is given by

$$
\begin{aligned}
p_2 &:= P(x \text{ executed after arrival}) \\
&= \sum_{k\geq 0, l\geq 0} \sum_{r\geq 0, q\geq 1} \int_0^\infty \gamma_B e^{-\gamma_B t} dt e^{-\gamma_B t} \frac{(\gamma_B t)^k}{k!} e^{-\gamma_S t} \frac{(\gamma_S t)^l}{l!} \\
&\quad \times P(M^B(k) < x < m^S(l), S(r,q), M^B(r) \vee S(q) \leq x < m^S(q-1)) \\
&= \sum_{r\geq 0, q\geq 1} \int_0^\infty dt \gamma_B e^{-\gamma_B t[2-F_B(x)]} e^{-\gamma_S t F_S(x)} P(S(r,q), M^B(r) \vee S(q) \leq x < m^S(q-1)) \\
&= \frac{\gamma_B}{\gamma_B(2-F_B(x)) + \gamma_S F_S(x)} \sum_{r\geq 0, q\geq 1} \binom{r-1+q}{r} p^r (1-p)^q F_B^r(x) F_S(x) [1-F_S(x)]^{q-1} \\
&= \frac{\gamma_B}{\gamma_B(2-F_B(x)) + \gamma_S F_S(x)} \sum_{q\geq 1} \frac{(1-p)^q F_S(x)[1-F_S(x)]^{q-1}}{[1-pF_B(x)]^q} \\
&= \frac{\gamma_B}{\gamma_B[2-F_B(x)] + \gamma_S F_S(x)} \frac{(1-p)F_S(x)}{p[1-F_B(x)] + (1-p)F_S(x)}
\end{aligned}
$$

For p_i^S the proof is similar.

Recall that A_1 is the event that buy order x is executed at arrival. Let us now compute $P(m^S_{\tau_x} \leq z | A_1)$ by computing

$$
\begin{aligned}
&P(m^S_{\tau_x} \leq z, A_1) \\
&= \mathbf{1}_{\{x\leq z\}} P(A_1) + \mathbf{1}_{\{x>z\}} \sum_{r\geq 0, q\geq 1} \int_0^\infty \gamma_B e^{-\gamma_B t} dt \int_0^z q(1-F_S(y))^{q-1} dF_S(y) e^{-\gamma_B t} \frac{(\gamma_B t)^r}{r!} e^{-\gamma_S t} \frac{(\gamma_S t)^q}{q!} F_B^r(y)
\end{aligned}
$$

$$= \mathbf{1}_{\{x \le z\}} p_1(x) + \mathbf{1}_{\{x > z\}} \int_0^z dF_S(y) \frac{\gamma_S \gamma_B}{[\gamma_B(2 - F_B(y)) + \gamma_S F_S(y)]^2}$$

Thus, we get

$$P(m^S_{\tau_x} \le z | A_1) = \mathbf{1}_{\{x \le z\}} + \mathbf{1}_{\{x > z\}} \frac{p_1(z)}{p_1(x)}$$

Therefore, the corresponding (conditional) density, denoted f^m_x, lives on $(0, x)$, and is given by

$$f^m_x(z) = \mathbf{1}_{\{x > z\}} \frac{p_1'(z)}{p_1(x)}$$

Proof is similar for the sell side.

Proofs for Section 3

Proof of Proposition 3.1: Similarly as with no machine, using the expressions for the distributions of M^B_t and m^S_t.

Proof of Proposition 3.2: We have

$$P(P_\tau \le x) = \sum_{r \ge 0, q \ge 1} P(S(r,q)) P(S(q) < M^B(r) \le \min\{x, s, m^S(q-1)\}, m^S(q-1) > b)$$

$$+ \sum_{r \ge 1, q \ge 0} P(B(r,q)) P(M^B(r-1) < s, M^B(r-1) \vee b \le m^S(q) \le B(r) \wedge x)$$

$$+ \sum_{r \ge 1, q \ge 0} P(B(r,q)) P(s \vee M^B(r-1) \le B(r) \le m^S(q) \wedge x, M^B(r-1) < s, m^S(q) > b)$$

$$+ \sum_{r \ge 0, q \ge 1} P(S(r,q)) P(M^B(r) \le S(q) \le b \wedge x, M^B(r) < s, m^S(q-1) > b)$$

The first term comes from the last order being a human sell order and trading with a human buy order in the book, and the second term from the last order being a human buy order and trading with a human sell order in the book. The third term comes from an incoming buy order trading with the machine, and the fourth term comes from an incoming sell order trading with the machine. The first two terms are computed similarly as with no machine. Conditioning on $B(r)$ in the third term and on $S(q)$ in the fourth, and computing the summations similarly as with no machine, we get the result.

■

Proof of Corollary 3.3: From the proposition, we have

$$P(P_\tau \le x) = \mathbf{1}_{\{x < b\}} \frac{p(1-p)}{(B-A)(D-C)} \int_C^x \frac{y - C}{[(1-p)\frac{b-C}{D-C} + p\frac{B-y}{B-A}]^2} dy$$

$$+ \mathbf{1}_{\{x > b\}} \frac{p(1-p)}{(B-A)(D-C)} \int_b^{x \wedge s} \frac{y - C}{[(1-p)\frac{y-C}{D-C} + p\frac{B-y}{B-A}]^2} dy$$

$$+\mathbf{1}_{\{x<s\}}\frac{p(1-p)}{(B-A)(D-C)}\int_b^x \frac{B-y}{[(1-p)\frac{y-C}{D-C}+p\frac{B-y}{B-A}]^2}dy$$

$$+\mathbf{1}_{\{x>s\}}\frac{p(1-p)}{(B-A)(D-C)}\int_s^{x\wedge B} \frac{B-y}{[(1-p)\frac{y-C}{D-C}+p\frac{B-s}{B-A}]^2}dy$$

$$+\mathbf{1}_{\{x>s\}}\frac{p}{B-A}\int_s^{x\wedge B} \frac{1}{(1-p)\frac{y-C}{D-C}+p\frac{B-s}{B-A}}dy$$

$$+\frac{1-p}{D-C}\int_C^{x\wedge b} \frac{1}{(1-p)\frac{b-C}{D-C}+p\frac{B-y}{B-A}}dy$$

We then get the density by differentiating.

Proof of Lemma 3.1: We have

$$E[N_b^2|\tau] = E\left[\sum_{n=1}^{\infty}\left(\sum_{i=0}^{n}\mathbf{1}_{b>S_i}\right)^2 e^{-\gamma_S\tau}\frac{(\gamma_S\tau)^n}{n!}\Big|\tau\right]$$

$$= \sum_{n=1}^{\infty}[np_S(1-p_S)+n^2p_S^2]e^{-\gamma_S\tau}\frac{(\gamma_S\tau)^n}{n!}$$

$$= \gamma_S\tau p_S+\gamma_S^2\tau^2p_S^2$$

After integrating over τ, we get

$$E[N_b^2] = \frac{1}{\lambda}\gamma_S p_S + \frac{2}{\lambda^2}\gamma_S^2 p_S^2$$

and analogously

$$E[N_s^2] = \frac{1}{\lambda}\gamma_B p_B + \frac{2}{\lambda^2}\gamma_B^2 p_B^2$$

Similarly, we have

$$\begin{aligned} & E[N_bN_s|\tau] \\ = & E[N_b|\tau]E[N_s|\tau] \\ = & \tau^2 p_S\gamma_S p_B\gamma_B \end{aligned}$$

so that

$$E[N_bN_s] = \frac{2}{\lambda^2}\gamma_S p_S\gamma_B p_B$$

The other expressions are proved in a similar fashion.

■

Proof of Proposition 3.3:

Since with uniform distribution we have

$$r_S = \frac{b^2-C^2}{2(D-C)}$$

$$r_B = \frac{B^2 - \hat{s}^2}{2(B-A)}$$

we need to maximize

$$\frac{1}{\lambda}\gamma_B[\frac{B^2 - s^2}{2(B-A)}] - \frac{1}{\lambda}\gamma_S\frac{b^2 - C^2}{2(D-C)}$$

$$-\rho\left[\frac{1}{\lambda}\gamma_B\frac{B-s}{B-A} + \frac{1}{\lambda}\gamma_S\frac{b-C}{D-C} + \frac{2}{\lambda^2}\left(\gamma_B\frac{B-s}{B-A} - \gamma_S\frac{b-C}{D-C}\right)^2\right]$$

The interior first order condition with respect to s is

$$s[1 + \rho\frac{4}{\lambda}\gamma_B\frac{1}{B-A}] = \rho + 4\rho\frac{\gamma_B}{\lambda}\frac{B}{B-A} - 4\rho\frac{\gamma_S}{\lambda}\frac{b-C}{(D-C)}$$

The interior first order condition with respect to b is

$$b[1 + \rho\frac{4}{\lambda}\gamma_S\frac{1}{D-C}] = -\rho + 4\rho\frac{\gamma_S}{\lambda}\frac{C}{D-C} + 4\rho\frac{\gamma_B}{\lambda}\frac{B-s}{(B-A)}$$

If $A = C$, $B = D$, $\gamma_B = \gamma_S = \gamma$, then, if we add the two conditions we get

$$(s+b)\alpha = \beta$$

where

$$\alpha = [1 + \rho\gamma\frac{8}{\lambda}\frac{1}{B-A}]$$

and

$$\beta = 8\rho\frac{\gamma}{\lambda}\frac{A+B}{B-A}$$

Subtracting we get

$$(s-b)\kappa = \delta$$

where

$$\kappa = 1$$

and

$$\delta = 2\rho$$

Solving this we get

$$s = \frac{\kappa\beta + \alpha\delta}{2\kappa\alpha}$$

$$b = \frac{\kappa\beta - \alpha\delta}{2\kappa\alpha}$$

and substituting we get

$$s = \frac{4\rho\gamma(A+B)}{\lambda(B-A) + 8\rho\gamma} + \rho$$

$$b = \frac{4\rho\gamma(A+B)}{\lambda(B-A) + 8\rho\gamma} - \rho$$

$\blacksquare$

TAB 9

THE IMPACT OF ALGORITHMIC AND HIGH FREQUENCY TRADING ON CME GROUP INC. MARKETS

Bryan Durkin
CFTC Technology Advisory Committee
July 14, 2010

CME Group Inc. appreciates the opportunity to present its views to the CFTC's Technology Advisory Committee. This paper discusses the role of algorithmic and high frequency traders at the exchanges that are owned by CME Group – Chicago Mercantile Exchange Inc., the Board of Trade of the City of Chicago, New York Mercantile Exchange, Inc. and the Commodity Exchange, Inc. (collectively, "CME Group markets"). In sum, we believe that (1) market participants who engage in high frequency or algorithmic trading contribute to CME Group markets by adding market liquidity, creating tighter markets, and in many cases providing continuous bids and offers in a market making capacity and (2) before any over-arching regulation is promulgated in that area, further study must be spent to understand the critical roles that such traders play in the futures markets (as well as the equity markets) and there should be a deep understanding of the existing self-regulatory systems, measures and controls which currently exist to monitor participants' activities to promote and ensure that those trading activities conform to DCM rules and regulations.

I. Introduction

a. Algorithmic/High Frequency Trading

While there is no precise definition of either term, algorithmic trading can be roughly defined as the use of any automated order execution methodology. CME Group views high frequency trading as a subset of algorithmic trading that is characterized by a trading system that systematically enters, often times smaller-sized, orders in an effort to minimize market impact. As noted by Cvitanic and Kirilenko in their paper, high frequency trading typically refers to trading activity that employs extremely fast automated programs for generating, routing, canceling and executing orders in electronic markets.

Algorithmic and high frequency trading represent the evolution of the markets from a floor-based model to an electronic model. In the traditional floor-based model, locals and other market-making entities played an important role in the price discovery mechanism. However, as spreads compressed due to multiple factors, including changes in minimum price increments, the profit opportunity of market making (liquidity providing) was reduced such that market makers needed to lower their costs in order to earn a return for the market risk that they assumed. To accomplish this, they turned to automation to reduce labor costs, increase operational efficiency and enhance risk management. As a result, algorithmic and high frequency trading have grown, contributing significant volume and providing greater liquidity and tighter bid/ask spread markets than were available in the traditional floor-based model. The liquidity generated by these traders is relied on by a variety of market participants including proprietary trading firms, investment banks, hedge funds and index traders.

Additionally, algorithmic trading is also used by both buy-side (institutional) and sell-side market participants to manage their order placement and execution process. Market participants have used automated order placement logic to reduce the market impact cost of larger orders. These users use strategies to systematically place orders into the market to achieve an execution benchmark, such as the volume weighted average price (commonly referred to as VWAP), over a defined time horizon. Market participants use automated order placement to reduce market impact and labor costs and increase operational efficiency.

It is also important to note that algorithmic and high frequency traders employ a wide variety of different strategies and do not act in concert. All of the trading roles and strategies performed by human intervention on the floor and on the screen, such as market-making, hedging, cash and derivative market arbitraging and order placement, are being replicated by algorithmic and high frequency trading. A significant proportion of high frequency traders active on the CME Group markets provide liquid markets by providing continuous markets in our products. Others engage in intra- and inter-market arbitrage opportunities by promoting trading among ETF's, futures and options as well as across stock index futures and cash equities.

Moreover, algorithmic and high frequency traders who provide liquidity take market risk; they have exposure to market movements notwithstanding the fact that their holding periods may be short in duration. They employ multi-factor models and other trading techniques in order to forecast "micro-prices" reflecting their view of current prices and movements. This participation contributes to price discovery and provides liquidity to other market participants.

In the futures markets, liquidity providers assume greater risk because the opposing side of every transaction is anonymous. This is not the case in other markets, such as certain secondary equity trading and cash options markets, as liquidity providers acting as market makers have a first look at the order flow and can internalize customer flow. Essentially, in those markets, liquidity providers acting as market makers are able to improve their quotes for certain market participants. In the futures markets, liquidity providers place orders and quotes into the markets which are then accessible to all other participants in the market. Liquidity providers in our markets have no informational advantage as they do not see customer orders prior to other market participants. To clarify, this does not imply that there is front running in other markets; rather, there is an informational advantage gained by knowledge of the frequency and distribution of incoming orders as it reflects market momentum that can be leveraged to improve trading by joining the momentum or trading correlated products in other markets. In addition, post trade, liquidity providers in our markets do not see the market participant on the other side of their trade.

II. Algorithmic and High Frequency Trading Provides Benefits to the CME Group Markets

With respect to the CME Group markets, algorithmic and high frequency trading plays an important role in the diverse mix of market participants who trade CME Group products, particularly when such trading is complemented by the types of exchange compliance and risk management procedures detailed in Section III.b below and strong risk management practices at the clearing firm level. Algorithmic and high frequency traders have evolved in response to advancements in technology and the growth of electronic trading. The percentage of trading volume attributable to algorithmic and high frequency trading will likely continue to increase in the future. As is discussed below, there is evidence that these traders are a significant component of trading activity in the CME Group markets and provide an important benefit in that they increase liquidity and transparency in the marketplace and narrow spreads, allowing investors to buy and sell our products at better prices and at lower costs. Algorithmic and high frequency traders engaged in trading simultaneously across multiple markets are able to bring these benefits to other markets.

CME Group has undertaken internal reviews to determine the impact of algorithmic and high frequency trading on the CME Group markets. The benefits deriving from algorithmic and high frequency trading that we have identified include:

- Providing liquidity
- Greater market depth (i.e., the number of contracts provided at the best bid-ask spread in the market)
- Increased price transparency
- Tighter bid/ask spreads
- Provision of continuous markets

In one review, we examined the link between the proportion of algorithmic trading activity and liquidity and volatility over a two-year period in a number of products, including, the e-mini S&P 500 future, EuroFX future, Eurodollar future, 10-year Treasury note future and crude oil futures (all nearby month except for 5th month Eurodollar). We reviewed the relationship between the percentage or proportion of daily volume attributed to algorithmic traders vs. the average bid/ask spread (market width) during regular trading hours.

Our results concluded that, in the majority of markets reviewed, algorithmic trading volume was positively correlated with a narrower bid/ask spread, increased depth of market and reduced volatility.

More recently, we reviewed activity in the nearby month e-mini S&P 500 futures for Q2 2010 to determine the type of activity engaged in by Automatic Trading Systems (ATSs) on the CME Globex system. Activity by ATSs is primarily conducted by entities engaged in algorithmic trading and high frequency trading.[1] One of the things that we examined was what proportion of the time an ATS acted as an "aggressor", i.e., was the ATS order the "lifting" side of a trade as opposed to was the order previously resting in the market. Based upon this review, we determined that, with respect to these markets, 54% of ATS volume was non-aggressor volume, i.e., 54% of ATS trades executed were orders where ATSs supplied liquidity to the market, as compared with the overall market as a whole, where 44% of the volume was characterized as non aggressor. In fact, ATSs provided nearly 60% of the liquidity in the market measured by their resting side volume share. This underscores the fact that these entities are a significant provider of liquidity, particularly when acting in a market-making capacity.

---

[1] The registration of ATS is further described in Section III.b.i below.

There has been speculation that the presence of high frequency traders contributed to recent market events but there is no evidence to support that contention. After the events of May 6, we undertook a review of high frequency trading activity to determine whether such trading caused or exacerbated the market events on that date. Based upon our review, it appears that the trading activity of a majority of the entities engaging in high frequency trading during the one-hour period surrounding the period of market stress were trades resulting from the firm's market making activities; the order activity and volume contributed by this user segment indicates that this group of participants did not pull out of the market during the market break and most continued providing liquidity in the futures markets during extreme market conditions. Based on our review, there is no evidence to support the proposition that high frequency trading exacerbated the volatility in the markets on May 6.

It is also important to note that automated trading or algorithmic trading originated and is used extensively in Europe. Accordingly, an unwarranted rush to place limits or impose regulatory burdens on high frequency trading in the United States may encourage such traders to shift the trading they currently conduct in the United States to Europe and other foreign jurisdictions that are already well-equipped to handle additional growth in both equities and futures.

In sum, based upon our own internal analysis, we believe that the use of algorithmic and high frequency trading by the variety of firms participating in the markets, including proprietary trading firms, investment banks, hedge funds and index traders, among others, has made the marketplace more efficient and competitive for all market participants. While the issue has not been extensively researched either within CME Group or elsewhere, careful consideration should be given to any decision to place significant restrictions or limitations on high frequency trading that would be harmful to the marketplace and result in less efficient and less liquid markets.

III. CME Has Effective Surveillance and Risk Management Tools in Place to Review Algorithmic Tradingand High Frequency Trading

In evaluating the potential role of regulation or other "best practices" as applied to algorithmic and high frequency trading, we believe that it is beneficial to assess those types of self-regulatory mechanisms and controls which the markets have developed in response to the rapid adoption of electronic trading over the past 10 years. Currently, CME Group markets are proactive in monitoring the trading activity of algorithmic and high frequency trading entities. In addition, CME Globex already employs many risk management policies and procedures which mitigate the risk associated with all types of electronic trading.

a. Market and Trade Practice Surveillance of Algorithmic and High Frequency Trading

As technology in financial markets continued to rapidly evolve over the past five years, algorithmic and high frequency trading has become increasingly significant in the context of order flow and transactions executed on CME Globex; order messaging, transactions and market data messaging have grown exponentially, now surpassing 5 billion messages per month. CME Group Market Regulation anticipated this explosive growth in messaging and designed systems infrastructure and applications to capture the full-range and contextual-support dimensions of all electronic activity, (e.g., all elements of order messages, transactions and market data quotations). This highly granular data includes identifying elements such as the Executing Firm, User Id and Account Number and whether the User is employing an automated trading system, as well as order details and numerous time stamps calibrated to the millisecond. Tools to efficiently mine and aggregate that data with exceptional speed are available to all regulatory analysts and allow for the querying of real-time and historical data. In addition, Market Regulation employs sophisticated systems to profile markets and participants, review and analyze participants' positions, generate live position and volume alerts based on anomalous activity, and identify transaction patterns and anomalies that may be indicative of misconduct.

Two of the primary systems used to monitor the activity of algorithmic and high frequency traders are SMART and RAPID. The SMART system contains detailed cleared trade data, quotation data, profile statistics of markets and market participants, analytical tools and a full suite of pattern detection capabilities, integrated with market and participant profiles, that allow analysts to set variable parameters and establish differential priority ranking for specific pattern elements. The pattern detection tools include identification of exceptions of potential trading abuses including, for example, cross trading, wash trading and money pass transactions. The RAPID system contains detailed CME Globex order and transaction data, market data and a suite of data mining applications for viewing, analyzing, summarizing and reconstructing market activity and states. It also allows for order book reconstruction, near real-time and historical reconstruction of participants' view of the displayed market, and is the source for the Live Alerting capabilities that have been developed.

The collection of systems employed by Market Regulation provides analysts with tremendous flexibility in analyzing market activity, including the activity of algorithmic and high frequency traders, at the most granular level.

Continuous market surveillance and administration is also performed by our Globex Control Center (GCC), which provides front-line customer service and market operation support for all electronic trading on CME Globex. The GCC uses a variety of both proprietary and 3rd party vendor tools to monitor real time trading. These tools include order entry alerts which monitor for potentially excessive order flow. When the order flow exceeds a predefined level for an individual connection (session), multi-tiered alerts are generated. Repeated rejected orders and database latencies are also monitored. The GCC is in the process of implementing executed volume alerts across the most active markets which provide reports on trading activity in markets based on volumes exceeding predetermined thresholds.

With respect to trade practice surveillance of algorithmic and high frequency trading, Market Regulation investigators monitor such activity using a variety of methods and approaches that leverage the technology described above. The primary sources from which algorithmic and high frequency trading-related investigations are derived are proactive internal research, referrals from the staff of the other CME Group departments, such as the GCC, or inquiries from external parties (e.g. firms, traders and customers). In addition to standard reviews of electronic trading activity for violations of rules related to, for example, cross trading and prearranged trading, wash trading, trading against customer orders, trading ahead and money passes, regular research is conducted of potentially manipulative or disruptive activity, and of anomalous messaging, transaction volumes, positions and prices movements that may be indicative of market misconduct. The Investigations area has also established an electronic trading team that identifies new strategies for examining potential trading abuses on the CME Globex platform.

b. Risk Management Policies and Procedures

In addition to the above-referenced surveillance procedures, CME Group has in place risk management processes, procedures and systems to preserve the integrity of its market in light of the many risks associated with maintaining a primarily electronic market, including that of algorithmic and high frequency trading. For example, the CME Group markets are the only exchanges in the world that require pre-execution credit controls which became mandatory for use by clearing members in June 2010.

i. Registration of ATS

All ATSs using CME Globex are required to identify themselves as an "ATS" and member ATSs or highly active ATSs must register.[2] This requirement was implemented at CME in 2006 and was expanded to include the NYMEX/COMEX markets in 2009. We currently have over 10,000 ATSs registered. ATSs are treated like any other market participant with respect to the aforementioned messaging policy. Additionally, the CME Group markets employ tools that allow for the monitoring of the trading activity of ATSs, as well as all other market participants, on both a real time and post-trade basis.

ii. Messaging Policies

CME Group has in place certain controls and policies which are designed to avoid problems associated with excessive messaging by market participants. CME Group has instituted a CME Group Messaging Policy that encourages market participants to trade and quote appropriately without harming market liquidity or system performance. Inefficient messaging slows system performance, negatively impacts other market participants and increases system capacity requirements and costs. To mitigate this, CME Group has implemented automated controls which monitor for excessive new order, order cancel and order cancel/replace messaging. If a session exceeds a designated message per second threshold over a three-second window, subsequent messaging will be rejected until the average message-per-session rate falls below this threshold. This helps CME Group to mitigate the extent of the impact of a malfunctioning trading system on the market.

CME Group has also instituted a policy of imposing fee surcharges for excessively high messaging rates. This policy benefits all customers trading on CME Globex by discouraging excessive messaging abuses, which in turn helps to ensure that CME Globex maintains the responsiveness and reliability of the system. Under the CME Globex Messaging Policy, clearing member firms are subject to the surcharges if they exceed product-specific benchmarks, individually tailored to the valid trading strategies of each market. CME Group calculates

[2] It should be noted that a trader who primarily enters orders manually but uses automated spreading activity is not considered an ATS and is not required to register.

benchmarks based on a per-product volume ratio, defined as the number of messages submitted for each executed contract in a given product.

In addition to the risk management policies implemented above, CME Globex employs a number of risk management functionalities which are designed to prevent potential problems which may result from electronic trading, including algorithmic and high frequency trading.

iii. Stop Logic Functionality

The CME Globex system has a Stop Logic functionality which serves to mitigate artificial market spikes that can occur because of the continuous triggering, election and trading of stop orders due to insufficient liquidity. If elected stop orders would result in execution prices that exceed pre-defined thresholds, the market automatically enters a brief reserved state for a predetermined time period, ranging from 5 – 20 seconds. During this period, no orders are matched but new orders other than market orders may be entered and orders may be modified and cancelled. The momentary pause that occurs when Stop Logic is triggered allows market participants the opportunity to provide liquidity and allows the market to regain equilibrium, thereby mitigating the potential for disruptive market moves.

iv. Price Banding Functionality

To ensure fair, stable and orderly markets, CME Globex subjects all orders to price verification using a process called price banding. The platform utilizes separate mechanisms for futures price banding and options price banding. Price banding prevents the entry of erroneous orders such as a bid at a price well above the market or an offer at prices well below the market which could trigger a sequence of market-moving trades that require subsequent cancellations. In order to determine the level of price banding, CME Group markets use the most current and relevant market information, including, for futures, trades, best bid and offer and implied bid and offer or indicative opening price, and for options, last price of an option or spread and a theoretical options price based on options pricing algorithms.

v. Protection Points for Market and Stop Orders

This CME Globex functionality automatically assigns a limit price (Protection Point) to futures market orders and stop orders to preclude the execution of these types of orders at extreme prices in situations where there is insufficient liquidity to support the execution of the order within an exchange-specified parameter of the current market.

The Protection Point values vary by product. The CME Globex system calculates the limit price for a Market Protected Order by applying the Protection Point value to the best bid or offer price (depending on the order's side of market) and by applying the Protection Point value to the trigger price for a Stop Protected Order. Any unmatched quantity remaining for a Market Protected or Stop Protected Order after it is executed to the Protection Point limit becomes a Limit Order at the Protection Point limit price.

vi. **Maximum Order Size Protection**

Maximum order size functionality on CME Globex prohibits entry of an order into the trading engine which exceeds a pre-determined quantity. This functionality provides protection against the so-called "fat finger" trades.

IV. Conclusion

CME Group believes that the use of algorithmic and high frequency trading by the variety of firms participating in the markets has made the marketplace more efficient, transparent and competitive for all market participants and will continue to do so as markets and technology eveolve. We believe that careful consideration should be given to any decision to impose restrictions or limitations on algorithmic and high frequency trading that would be harmful to the marketplace and result in less efficient and less liquid markets. However, CME Group believes that algorithmic and high frequency trading must be complemented by the types of exchange compliance and risk management procedures described herein and by strong risk management practices at the firm level.

Consequently, we believe that (1) market participants who engage in high frequency or algorithmic trading contribute to CME Group markets by adding market liquidity, tighter markets, and in many cases provide continuous bids and offers in a market making capacity and (2) before any over-arching regulation is promulgated in that area, further study must be spent to understand the critical roles that such traders play in the futures markets (as well as the equity markets) and there should be a deep understanding of the existing self-regulatory systems, measures and controls which currently exist to monitor participants' activities to promote and ensure that those trading activities conform to DCM rules and regulations.

TAB 10

Market Access
Risk Management Recommendations

April 2010

On behalf of the Futures Industry Association Market Access Working Group, we are pleased to present recommendations for managing the risk of direct access trading. Recognizing the importance of promoting best practices in this area, the FIA board of directors in January 2010 agreed to assemble a committee to formulate best practices for direct access to exchanges. The group includes representatives from clearing firms, trading firms, and exchanges. The scope of their work includes pre-trade order checks, post-trade checks, co-location policies, conformance testing, and error trade policies.

The study will be shared with futures and options exchanges around the world. Later this year, FIA plans to survey exchanges that offer direct access to determine what types of risk controls are in place and publish the results of the survey.

We appreciate the time and resources the members of the Market Access Working Group contributed to the creation of this document. This is not the first group FIA has convened to address risk management practices. In 2004, FIA published a series of recommendations on error trade polices. In 2007, FIA published the results of a survey on risk controls at key exchanges. In 2009, the FIA/FOA Clearing Risk Study included recommendations for pre- and post-trade risk controls.

We expect the need for risk controls to continue to evolve as the industry evolves and FIA is committed to monitoring and supporting practices and procedures that improve the integrity of the markets.

Yours truly,

Peter Johnson
Chairman
Market Access Working Group

FIA Market Access Working Group

The following organizations participated in the development of the FIA Market Access Risk Management Recommendations*:*

- Bank of America Merrill Lynch
- Barclays Capital
- CME Group
- Credit Suisse
- DRW Trading
- Eurex
- Geneva Trading
- IntercontinentalExchange
- J.P. Morgan Futures
- Newedge Group
- Nico Trading
- NYSE Liffe
- XR Trading

The FIA is the U.S.-based international trade association which acts as a principal spokesman for the futures and options industry. Its membership includes the world's largest futures brokers as well as derivatives exchanges from more than 20 countries.

Introduction

Managing the risk of providing direct access to an exchange's network is a critically important responsibility of all parties involved in the process—clearing firms, exchanges, and the direct access firms themselves. However, managing such risk must be done in a manner that does not disadvantage one direct access firm over another solely because it, or its clearing firm, endeavors to act more responsibly. This can only be done if exchanges themselves provide basic risk management tools, and construct them in such a manner that latency is identical to all direct access firms, no matter how clearing firms utilize such tools. Indeed, this will encourage the clearing firm to employ such tools in the most responsible fashion, without fear that it will lose business to other clearing firms that do not act so responsibly.

Recognizing the importance of promoting best practices in risk management of direct access trading, the FIA board of directors in January 2010 established a Market Access Working Group to identify risk-specific controls that are already in place at exchanges, clearing and trading firms and recommend controls that should be in place as a matter of best practice before allowing direct access. The MAWG consists of representatives from clearing firms, exchanges, and trading firms. The group has been meeting since January to agree on recommendations for pre- and post-trade risk controls, co-location, conformance testing, and error trade policies.

Latency-sensitive traders, which rely on direct access, can play a vital role in the marketplace, bringing liquidity to the markets, reducing volatility, tightening bid-ask spreads, and contributing to price discovery[1]. The recommendations presented here represent another step in improving the way direct access risk is managed. The industry has been working together for several years to ensure risk management practices reflect the realities of the current trading environment. In 2004, FIA published a series of recommendations with respect to exchange error trade policies and procedures. In 2007, FIA published a "Profile of Exchange and FCM Risk Management Practices for Direct Access Customers," which identified issues with this type of trading and enumerated the results of a survey of risk controls at key exchanges. The FIA/FOA Clearing Risk Study, released in February 2009, included recommendations for exchanges to implement pre-defined authorizations, position limits, and monitoring and intervention capabilities.

The current project establishes principles the industry should consider when allowing direct access to exchanges. Although the guidelines contained in this document are more generally suited to futures and options markets, many of the principles and recommended implementations are applicable to other types of markets. The MAWG recognizes that market structures vary and exchanges need to implement risk controls across multiple product lines. For example, some exchanges offer both equities and futures on the same trading platform. The MAWG also acknowledges that exchanges are in varying stages of permitting direct access and therefore these recommendations may not be immediately achievable. Instead, these recommendations are put forth as agreed-upon principles that the global futures industry needs to work toward implementing. In addition, the MAWG recognizes that these recommendations must be considered in the context of the regulatory structures in which markets operate.

[1] See *Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market* by Alain Chaboud, Benjamin Chiquoine, Erik Hjalmarsson, Clara Vega, in which the empirical data examined by the authors suggested that, in the spot interdealer foreign exchange market, "the presence of algorithmic trading reduces volatility" and "computers *do* provide liquidity during periods of market stress." (International Finance Discussion Paper, Board of Governors of the Federal Reserve System, dated October 2009, p. 26.)

This document is designed to serve as a framework for developing risk controls. It attempts to strike the right balance between guiding principles and prescriptive mandates. Accordingly, this document reflects two types of recommendations: principles and implementation recommendations. The first type is a guiding principle that describes the type of control and what should be achieved by implementing the control. The principles, in some cases, are followed by implementation recommendations.

The document includes a section on co-location and proximity hosting. Co-location and proximity hosting have often been included in discussions related to risks associated with high-frequency trading, but the MAWG does not believe this is a risk management issue. Co-location and proximity hosting refer to data centers that offer an alternative method to brokerage and trading firms seeking the fastest possible access to an exchange's network and are not inherently risky. Co-location takes place when the exchange provides connectivity and hosting in its own data center via its own network. Proximity sites are data centers offered by an exchange or a third-party vendor for low-latency access to an exchange's network via a third-party network connection.

Background

Direct access firms either join the exchanges as non-clearing members (NCMs) or access the exchanges in the name of their clearing member. While there is no distinction between a direct access firm that becomes a non-clearing member of an exchange and one that does not when it comes to risk and credit controls, NCMs are subject to an exchange membership approval and vetting process. NCMs also are subject to exchange rules such as market manipulation, wash trades and message limit violations. In either case, these firms' transactions must be financially guaranteed by a clearing member before the exchange grants direct access to these firms. The clearing firm guarantees the trades pursuant to an agreement with the trading firm and retains administrative and risk control over orders submitted to the exchange trading engine.

There are three ways a non-clearing firm can access the exchange network directly:

a. Direct access via a clearing firm (DA-C)—trading firm orders pass through the clearing member's system prior to reaching the exchange trading engine.
b. Direct access via vendor (DA-V)—trading firm routes orders through a vendor controlled by the clearing firm or other third-party infrastructure to the exchange trading engine.
c. Direct access to the exchange (DA-E)—trading firm routes orders directly to the exchange trading engine without passing through the clearing member or a third-party infrastructure.

Risk management of direct access market participants is not the exclusive responsibility of exchanges, clearing firms or even the direct access firms themselves. Rather, exchanges, clearing firms, and direct access firms each have a role in ensuring that appropriate risk controls are in place for this type of market access. Clearing firms that frequently manage many exchange interfaces would benefit greatly from standardization of risk management controls across exchanges. The more standardization of risk controls, the more efficiently and effectively clearing firms are able to monitor and manage the risks associated with direct access clients.

Role of Direct Access Participant

Trading firms typically have risk controls in place to monitor and risk-manage their trading systems. These protections operate within their risk model and include pre-trade risk controls e.g. order size limits. Below is a sample of risk controls frequently employed by trading firms. Although these controls represent good practice, they are not uniformly enforceable by exchanges or clearing firms.

- Conformance Testing. Trading firms are required to pass conformance testing with the party providing access when implementing a new direct access system or when the exchange deems it necessary because of a fundamental change in functionality on the exchange side. The onus is on the trading firm to determine when it must recertify due to a change in logic within its system.
- Heartbeating with the Exchange. Trading systems can monitor "heartbeats" with the exchange to identify when connectivity to the exchange is lost. If connectivity is lost, the system is disabled and working orders are cancelled.
- Kill Button. Trading systems can have a manual "kill button" that, when activated, disables the system's ability to trade and cancels all resting orders.
- Pre-Trade Risk Limits. Trading firms can establish and automatically enforce pre-trade risk limits that are appropriate for the firms' capital base, clearing arrangements, trading style, experience, and risk tolerance. These risk limits can include a variety of hard limits, such as position size and order size. Depending on the trading strategy, these limits may be set at several levels of aggregation. These risk limits can be implemented in multiple independent pre-trade components of a trading system.
- Post-Trade Risk Limits. Trading firms can also establish and automatically enforce post-trade risk limits that are appropriate for the firm's capital base, clearing arrangements, trading style, experience, and risk tolerance. For example, a trading firm can set daily loss-limits by instrument, asset class, and strategy and automatically close out or reduce positions if those limits are breached.
- Fat-Finger Quantity Limits. Trading systems can have upper limits on the size of the orders they can send, configurable by product. They can prevent any order for a quantity larger than the fat-finger limit from leaving the system.
- Repeated Automated Execution Throttle. Automated trading systems can have functionality in place that monitors the number of times a strategy is filled and then re-enters the market without human intervention. After a configurable number of repeated executions the system will be disabled until a human re-enables it.
- Near-Time Reconciliation. Trading systems can have functionality in place that accepts drop-copies from exchanges and clearing firms. Drop copies are duplicate copies of orders that allow a firm to compare the exchange or clearing firm view of trades and positions with the firm's internal view. This helps to assure that all systems are performing as expected and maintaining accurate and consistent views of trades and positions.
- Reasonability Checks. Trading systems can have "reasonability checks" on incoming market data as well as on generated values.

Role of Clearing Firm

The management of client risk by clearing firms, and of clearing member risk by clearinghouses, has evolved as trading has moved from exchange floors to computer screens. In most respects, risk controls have strengthened.

Clearing firms direct significant resources toward managing and monitoring risk and refining approaches to assessing clients' risk exposure. Clearing firms frequently employ the following risk management controls with direct access clients:

- Most exchanges and self-regulatory organizations (SROs) require the clearing firm to ensure that the trading firm has pre-trade risk controls in place. Clearing firms may require the trading firm to provide network access to the trading firm's pre-trade risk controls to allow a clearing firm to set various risk limits and, if appropriate, stop the trading firm's trading. Network access is technically difficult to achieve, however, and trading firms can override risk controls set by clearing firms.
- The clearing firm will conduct substantial due diligence on prospective direct access clients and will grant direct access rights only to those clients who are deemed sufficiently creditworthy and whose internal controls are deemed sufficiently strong that pre-trade monitoring by the clearing firm is less essential. A clearing firm may also require additional collateral to provide further certainty that the trading firm will be able to meet any obligations that might arise from trading. In addition, the clearing firm will monitor the trading firm's account to determine whether margin requirements are being met.
- Trading firms are judged on their willingness to share information with their clearing firm. The more transparent a client is willing to be, the more likely the clearing firm is to grant direct access.
- Clearing firms have risk controls built into order entry systems they offer trading firms. These risk controls include many of the controls described later in this document.
- Increasingly, clearing firms are depending on the exchanges to provide pre-trade risk controls. Often, limits on the exchange systems can be configured and monitored by the clearing firms. This ensures that risk controls do not become a source of competition between clearing firms.
- Finally, clearing firms have agreements with trading firms that require the trading firms to have specified risk controls in place, restrict access to authorized personnel, and comply with relevant rules. Clearing firms monitor and enforce compliance with these agreements on an ongoing basis.

Role of Exchange

The primary business and function of exchanges is matching and clearing trades, regulating their market, and ensuring that the market operates safely with minimal systemic risk in order to sustain the overall viability of the market. The default or failure of the client of a clearing member has no immediate risk consequences for the clearinghouse unless it causes losses that lead to the default or failure of the clearing member. However, the provision of controls to help avoid such events must be regarded as a priority of any exchange in order to protect the overall integrity of its marketplace, and in recognition and support of the risk management role undertaken by clearing members.

Exchanges have in place well-defined policies and procedures describing the responsibilities of clearing firms and direct access firms.

- Exchange rules may require that clearing firms implement specified risk management standards with regard to direct access clients. The exchange's requirements and onboarding processes for clearing firms and their direct access customers encompass and support the risk management standards. The exchange processes may include: legal paperwork, system certifications, and permissioning security.
- Clearing firms for directly connected entities must follow recommended exchange guidelines for direct access, including in many cases requirements that clearing firms configure and monitor automatic risk limits and that they maintain the ability to halt a client's trading system, if appropriate.
- Exchanges have the ability to establish an error trade policy that provides a uniform set of policies and procedures that are followed in the event of an error.
- Exchanges have the ability to enable or restrict access per established rules.
- Exchanges establish rules surrounding processes to ensure that direct connections are guaranteed by clearing firms.
- Exchanges make non-clearing entities and system providers aware of exchange rules and responsibilities in the processes surrounding connectivity and electronic trading and ask them to certify to the exchange and clearing firm their capabilities to provide risk management functionality.

1. Execution Risk Tools

Pre-trade order checks are risk controls put in place to prevent execution of a trade because of error or "fat-finger" problems, or a client trading beyond authorized trading limits. Pre-trade risk controls can be put in place at the trading firm, clearing firm, or exchange level. Pre-trade risk controls have become a point of negotiation between trading firms and clearing members because they can add latency to a trade. To avoid such negotiations, the MAWG believes that certain risk controls should reside at the exchange level and be required for all trading to ensure a level playing field. The right to set and manage, or authorize a trading firm to set and manage, any pre- or post-trade order checks at the exchange's matching engine, however, should reside with the clearing firm.

Recommended Implementation:

- To reduce the inevitable errors that occur with manual data entry, exchanges should work towards providing a standard communication protocol that would allow firms to automate setting and updating risk parameters for individual trading entities. This would also give clearing firm risk managers the ability to more efficiently disable a client from multiple exchanges simultaneously. An API based on an agreed standard protocol such as FIX would be the preferred method for entering and updating limits.
- Unless otherwise indicated, exchange risk control systems should provide clearing firms with the ability to define risk controls by product. All limits should be set by positive permissioning. The auto-default should be set to zero (i.e. clearing firm will set limits only for the products that they are allowing the trading firm to trade).

a. Order Size

Quantity-per-order limits are the most basic types of pre-trade risk management tools to help prevent accidental "fat-finger" incidents. This type of limit sets a maximum number of contracts that can be bought or sold per order.

Principle:

Quantity-per-order limits should be mandatory:
(a) The clearing firm should establish limits with the trading firm to avoid generating and sending erroneously-sized orders to the market. Occasionally, larger-sized orders are legitimate. In such cases, the trading firm needs to contact the clearing firm to adjust their limits.
(b) The exchange should provide default limits to protect the integrity of its market.

Recommended Implementation:

A clearing firm providing direct access to a market should have visibility to the limits and the ability to set appropriate limits for the trading firm's activity, regardless of whether the trading firm accesses the market directly (DA-E), through the clearing member system (DA-C) or through a third-party system (DA-V).

- Risk controls need to be sophisticated enough to allow the clearing firm to set pre-trade limits per product for each client and prevent trading beyond established limits. Different sized limits are required for more liquid versus less liquid instruments (e.g., front month versus back month futures or options, in-the-money versus out-of-the-money options).
- Trading firm access to products should be blocked until limits are established by the clearing firm. Default limits should not allow "unlimited" trading. In addition, the clearing firm would like to have the ability to set controls for multiple products at one time.

b. Intraday Position Limits

Intraday position limits give the clearing firm the ability to block a trading firm from increasing its positions beyond a set threshold. Limits placed at the exchange level, rather than the order-entry system, allow centralization and standardization of risk controls. Position limits, however, are intended as "speed bumps on trading" and not as actual credit controls. These limits include start-of-day positions, cash in account, and cross-asset margining. Position limits provide the ability to automatically halt errant algorithms before credit limits are exceeded. Once a trader is blocked, the risk department has time to perform a risk evaluation before allowing further trading.

Principle:
The exchange should make available the ability to set pre-trade intraday position limits. Once the trading entity has reached these limits, only risk-reducing trades would be allowed.

Recommended Implementation for Futures:
The position limit capability should have the following characteristics:

- Set by trader, account, or firm and with the ability to set by groups of traders or accounts.
- Set maximum cumulative long positions and maximum cumulative short positions.
- Include working orders in maximum long/maximum short position calculations.
- Set by product level.
- Provide the ability to raise or lower limits intraday.
- Be configurable by open API, preferably FIX API.
- Be mandatory for all participants so that latency is the same for all.

Recommended Implementation for Options:

- Recognizing that options have a lower delta than futures, position limit capability must include the ability to differentiate limits by product type.

c. Cancel-On-Disconnect

When a system unintentionally disconnects from the exchange network, it creates uncertainty about the status of working orders. Automatic cancellation of orders upon disconnect provides certainty to the trading firm and risk manager whether orders have been filled or cancelled. Some users, however, may not want to have their orders automatically pulled from a market as the working order may be part of a hedged position or a cross-exchange strategy trade.

Principle:
Exchanges should implement a flexible system that allows a user to determine whether their orders should be left in the market upon disconnection. This should only be implemented if the clearing firm's risk manager has the ability to cancel working orders for the trader if the trading system is disconnected. The exchange should establish a policy whether the default setting for all market participants should be to maintain or cancel all working orders.

d. Kill Button

A "kill" button provides clearing firms with a fast and efficient way to halt trading activity at the exchange level when a trading firm breaches its obligations vis-a-vis the clearer (e.g. by exceeding credit limits due to erroneous activity of an automated trading application). The trading firm will be excluded from trading until the clearing firm explicitly reinstates it.

Principle:
Exchanges should provide clearing firms with the ability to: 1) delete all open orders and quotes and 2) reject entry of new orders and quotes.

Recommended Implementation:

- The exchange should have a registration system that requires firms to specify which staff members are authorized to use the kill button.
- The system itself should have explicit warnings informing authorized users of the consequences of activating the kill button.
- Similar functionality could be implemented to allow a trading firm to halt trading activity on a firm-wide, trading group or individual trader basis.

e. Order Cancel Capabilities

Principle:
Exchanges should provide to clearing members an order management tool that allows real-time access to information on working and filled electronic orders. The tool should provide risk mitigation functionality in the event of an electronic trading system failure.

Recommended Implementation:
The clearing member and trading firm should have the ability to view and cancel orders via this tool. Clearing members should be able to delegate and permission the tool for individual traders or firms at granular levels.

The tool should provide view capabilities for:

- current order status
- fill information, including partial fills
- cancel and replace history
- order timestamps

The tool should provide cancel capabilities for:

- individual orders
- groups of orders
- all working orders via a single command

f. Price Banding/Dynamic Price Limits

Price banding or dynamic price limits are an automated order-entry screening process designed to prevent entry of buy or sell orders priced substantially through the contra side of the market. It reduces the number of error trades that take place in the market by preventing bids from being entered too far above current market prices and offers from being entered too far below current market prices.

Principle:
The exchange should have the ability to set price limits on a dynamic basis, continuously adjusting throughout the day to account for current market conditions.

Recommended Implementation:
Exchanges should have the ability to widen price bands throughout the trading day when necessary to account for additional volatility in the market. The width of the price limits should be determined by product. Price banding occasionally can be too strict for less liquid markets and may need manual intervention to facilitate trading if the current range is deemed unsuitable.

Price banding for options requires a different approach because options are more dynamic. Price banding may be too restrictive for less liquid options contracts because of wider bid-ask spreads.

g. Market Maker/Sweep Protections

Sweep protections are designed for firms with specific market-marketing obligations to quote options en masse. Although these protections are most frequently used in options markets, they can be applied to other markets. Market-maker protections are parameters set by market makers and implemented by the exchange to provide a degree of risk protection by limiting the market maker's quote execution exposure.

Principle:
Exchanges should allow a level of protection for market makers who quote simultaneously on both sides of the market.

Recommended Implementation:
Protection parameters should be optional and should allow values to be set by each market maker or market-making entity. When market maker-defined protection values are met or exceeded within certain time intervals, the protections should be triggered. When triggered, the electronic trading system would initiate the market-maker protection functionality, which rejects new messages and/or cancels resting quotes associated with the market maker.

h. Internal Trade Crossing

It is common for multiple independent trading strategies to be active at the same time within a single firm. The strategies may interact on the market by taking opposite sides, occasionally generating inadvertent wash trades. This is a common situation with direct access and the increasing use of broker execution algorithms that may stretch orders over a period of time, micro-manage slices that may interact with another order placed by the same legal entity, or run as an auto-hedging facility with no intention upfront to create a wash trade.

The MAWG considered whether technology could assist risk managers in identifying wash trades. The group concluded that it is impossible for exchanges to implement such risk controls because account ownership information is not available at the matching engine. While clearing members have the ownership information and can confirm whether a client resides in the same profit center of the firm, algorithms may be producing orders that interact with accounts within the same legal entity. Further, customers can use multiple systems within a legal entity that don't necessarily interact with each other on a pre-trade basis. The MAWG concluded that there was no way to design a rule that would prevent wash trades without preventing legitimate trades.

Principle:
Wash trades are prohibited to prevent manipulating the market by artificially distorting market price or volume. Inadvertent crosses do not have the intent to mislead the public. Exchanges, working within the framework provided by their respective regulators, should set guidelines for vendors, customers, and clearing members, defining what would be acceptable reasons for inadvertent cross trades. Existing rules should be re-examined in the context of today's trading environment.

2. Post-Trade Checks

In addition to pre-trade risk controls, post-trade checks allow clearing and trading firm risk managers to track all working/open orders and all executed and cleared orders. "Drop copy" functionality gives clearing firms the ability to monitor orders on a near real-time basis without adding latency to the order flow. Drop-copy functionality allows clearing members to receive duplicate copies of client working/executed orders as they enter the exchange network and/or are matched at the clearinghouse.

Principle:
Exchanges should make drop copies available to clearing and trading firms.

- Trade capture drop copy: Exchanges should provide clearing firms with drop copies of orders and executed trades. This allows clearing firms to get their current set of trades and positions from a secondary channel independent of the primary trading system.
- Post-clearing drop copy: Exchanges should provide clearing firms net position per maturity per contract as soon as the trade is matched at the clearinghouse. This functionality needs to be as close to real-time as possible.
- Exchange drop-copy functionality should allow clearing firms to enable trading firms to receive trade capture and post-clearing drop copies.

Recommended Implementation:
The post-clearing drop copy feed should contain all messages including acknowledgements, fills, amendments and cancellations. Exchanges need to work toward an industry standard of delivering cleared information in a maximum of two-three minutes after a trade is executed. This data needs to be delivered via a standard protocol, preferably via FIX API.

3. Co-Location Policies

When considering co-location, exchanges should recognize that one of the main benefits of such a service is that it creates a level playing field for firms that want low-latency access to the exchange. It provides firms, both large and small, with low-latency connectivity for a reasonable cost made possible by the exchange sharing the costs of the required technical infrastructure with interested participants. When co-location and proximity sites are not available, it encourages firms to seek confidential knowledge about matching engine locations and compete for building space closest to those engines so they can build their own private data centers. This exacerbates the differences in the ability of market participants to obtain market access.

Principle:
Steps should be taken to ensure that access to co-location is available to every firm that is interested in such a service and that the terms of the co-location service remain transparent to all market participants.

4. Conformance/Certification Testing

Principle:

- All trading firms that wish to write directly to the order entry or market data interfaces of an exchange should be required to pass an initial set of conformance tests for execution and market data that highlight basic functionality of the trading system that will be making the direct connection. All ISVs and proprietary systems should be required to pass the same conformance tests, so the proprietary system client using the ISV should not be required to pass conformance.
- The exchange should be required to provide a conformance environment on-demand for re-certification requirements.

Recommended Implementation:

A representative of the exchange should interview the proprietary system client to determine which functionality should be tested. Exchanges should test the ability of a direct access firm to:

- Send a request for and process the exchange's response for the following: Log On, Log Off, New Order, Cancel, Order Modify, Sequence Reset, Instrument Definition Requests, and Market Snapshot requests.
- Process the following exchange messages: Business Reject, Session Reject, Complete Fills, Partial Fills, Exchange Open/Close, Market Data Updates, Trade Updates.
- Properly handle the exchange recovery mechanism provided when messages are sent from the exchange to a proprietary system participant, but the client isn't actively connected.
- Recertification should be required whenever core functionality has changed at the exchange. It should be up to the exchange to decide what functionality needs to be recertified as well as to notify each proprietary system participant of the need to recertify.
- Recertification should be required whenever a participant's core functionality has changed. It is up to the proprietary system participant to notify the exchange when this happens as well as to schedule the conformance test.

5. Error Trade Policy

The potential for trading errors by direct access traders causing significant market disruptions is of utmost concern to all market participants and regulators. Although traders and trading system engineers have an incentive to build robust systems and safeguards to avoid potential error trade situations and the substantial costs associated with them, the potential for error trades still exists. Robust pre-trade risk controls such as price banding significantly reduce the potential for erroneous trades but exchanges still need to enforce a strict error trade policy.

A robust error trade policy minimizes systemic risk by affording market participants confidence that when an error trade occurs, it will be evaluated and resolved according to a uniform set of policies and procedures. Conversely, subjectivity or ambiguity in an error trade policy amplifies risk through uncertainty. The objective of an error trade policy should be to remove the uncertainty of open-ended market exposure and allow traders to expeditiously resume normal trading activity. This is critical for maintaining market confidence and continuity.

a. Trade Certainty

An important aspect of market integrity is the confidence that, once executed, transactions will stand and will not be subject to arbitrary cancellation.

Principle:
Exchanges should adopt a "Preferred Adjust-Only Policy" to ensure absolute trade certainty to all parties to an error trade. In a Preferred Adjust-Only Policy all trades inside of a product-specific "no-adjust" range are ineligible for adjustment. All trades outside of the no-adjust range potentially could be adjusted to the edge of the no-adjust range from the prevailing market at the time of execution. The Preferred Adjust-Only Policy would not eliminate the authority of an exchange to cancel or correct trades under extreme circumstances.

b. Contingency Orders

The most challenging aspect of an error trade policy is the appropriate way to handle a contingency or stop order triggered by an erroneous transaction. The MAWG recognizes that a clearing firm could incur losses on contingency orders their customers placed which were filled as the result of an erroneous trade but cannot be passed on to the customer since the adjusted price does not indicate that the order should have been filled.

Principle:
In keeping with the objective of the Preferred Adjust-Only Policy, contingent or stop orders executed as a result of an error trade should be eligible for compensation from the party that made the error. An exchange's authority to cancel orders under extreme circumstances should not be invoked merely because an order is a contingent order.

c. Notification

Markets continue to trade while the parties to a trade and the exchange determine whether a trade is erroneous. The identification of a possibly erroneous trade well after it has been executed and its later cancellation can create even more uncertainty in the market. Market integrity, therefore, demands that exchange policies and procedures establish strict, narrow time frames in which a request to cancel a trade is made.

Principle:

The exchange should establish a minimal reporting time of less than five minutes for firms to notify the exchange that an error has occurred.

The exchange should announce a potential adjust-or-bust situation immediately upon notification and the adjust decision should be disseminated to the marketplace within a reasonable timeframe via a specific market data message, email and/or other established mode of communication on a best efforts basis.

Futures Industry Association
2001 Pennsylvania Avenue N.W., Suite 600
Washington, DC 20006-1823
202.466.5460 • www.futuresindustry.org