



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

Alain Chaboud, Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega

1

October, 2012

Our Papar

Our Paper

- Novel dataset
 - Algorithmic trading in the foreign exchange market: euro-dollar, yendollar, euro-yen
 - September 2003 to December 2007
 - Observe four types of trades: HH, HC, CH, CC
- New version:
 - Novel way of measuring the correlation of algorithmic trading actions
 - **Focus is on the effect AT has on price discovery**

Our Paper (continued)

- Role AT plays in price discovery process
 - AT measured in five different ways
 - AT participation
 - AT liquidity provision
 - AT liquidity demand
 - AT "signed" liquidity demand
 - Correlation of AT trading actions
 - Price efficiency:
 - Triangular arbitrage opportunities
 - Autocorrelation of high frequency returns
- Biggest challenge: endogeneity
 - Heteroskedasticity identification approach (Rigobon (2003), Rigobon and Sack (2003, 2004))



Theory

- Disagreement on the effect AT may have on price discovery (Foucault (2012))
- Positive view: Biais, Foucault, and Moinas (2011) and Martinez and Rosu (2011)
 - Computers are fast and better informed than other traders
 - Computers use market orders to exploit their informational advantage
 - Computers make prices more informationally efficient, but increase adverse selection costs for slow traders

Theory (continued)

- Negative view: Jarrow and Protter (2011): computers reacting to a common signal, create price momentum and push prices further away from fundamentals
- Negative view: "crowding effect": Kozhan and Wah Tham (2012) and Stein (2009) computers entering the same trade at the same time pushes prices further away from fundamentals
- Positive view: Oehmke (2009) and Kondor (2009), competition among convergence traders makes prices more informationally efficient

Theory (continued)

- Negative view: If computers are "noise" traders: Delong et al. (1990), Froot, Scharfstein, and Stein (1992)
 - Positive feedback traders who predictably extrapolate past price trends
 - Short-term speculators (chartist) herd and put too much emphasis on some (short-term) information and not enough on fundamentals
 - AT could cause "excessive" volatility
- Foucault (2012) effect may depend on strategy computers specialize on.

What is Algorithmic Trading?

- Algorithmic Trading (AT): The use of computer algorithms to manage the trading process. Formulate trading decisions and execute trades.
- In practice: Automated execution, computers directly interact with electronic trading platforms. Very fast.
- It includes High Frequency Trading and other types of algorithmic trading
- First AT trade on EBS in 2003. Fast growth on EBS



Participation Rates of Algorithmic Traders





- EBS (essentially the global site of price discovery in interdealer FX market for several large currencies) records when a trade is placed manually (keyboard) or by a computer interface
- Minute-by-minute data from 2003 to 2007
- Three currency pairs (EUR-USD, USD-JPY, EUR-JPY)



Our Data (continued)

- Volume and direction of trade breakdown each minute by AT (Computer) and non-AT (Human).
 - We know how much computers "take" from human "makers."
- Four possible types of transactions: HH, HC, CH, CC (maker-taker).



Trading Volume by Maker and Taker Type

Classifying Trades

- · Maker: Trader who posted price.
- Taker: Trader who bought or sold at posted price.
- · More makers: higher market liquidity.

Maker Type	Taker Type	
Human	Human	(HH)
Computer	Computer	(CC)
Human	Computer	(HC)
Computer	Human	(CH)







Our Data (continued)

Five different measures of AT activity

- AT participation:
 Vol(CH+HC+CC)/Vol(CH+HC+CC+HH)
- AT liquidity supply:

Vol(CH+CC)/Vol(CH+HC+CC+HH)

AT liquidity demand:

Vol(HC+CC)/Vol(CH+HC+CC+HH)

AT signed liquidity demand:

|OF(HC+CC)|/(|OF(HC+CC)|+|OF(CH+HH)|)

Correlation of AT trading actions: R-measure

What if algorithmic traders (ATs) all did the same trade at the same time?

- Correlated strategies can make prices more informationally efficient ("convergence" trades)
- Correlated strategies can cause excess volatility
- Yen-Dollar market on August 16, 2007





Human-Maker / Computer-Taker Order Flow Yen/\$ _ 117 Computer-Maker / Computer-Taker Order Flow \$ Millions 1500 _T \$ Millions 1500 ₇ Order Flow ----- Dollar-Yen 1000 1000 116 500 500 115 0 ______ 0 --500 114 -500 -1000 -1000 113 W -1500 -1500 112 -2000 -2000 111 -2500 -2500 12 AM 6 6 12 6 12 6 12 PM AM PM PM AM AM PM









Do algorithmic trades, strategies, tend to be correlated?

- We do not know strategies, we do not have orders, only completed trades.
- Instead: Do computers trade with each other as much as expected, as much as random matching would predict? If computer strategies are correlated, we should observe less trading among computers than expected.
- More precisely: Do computers "take" from humans and computers in the same proportion as humans take from humans and computers?

- Prob(HC)/Prob(CC) = computer taker ratio = RC
- Prob(HH)/ Prob(CH) = human taker ratio = RH
- In a world with more human makers than computer makers (our world), we expect Prob(HC)/Prob(CC) > 1, i.e., computers take more from humans than from other computers. And we expect Prob(HH)/ Prob(CH) > 1, i.e., humans take more from humans than from computers.
- However we expect RC/RH=1, i.e., humans take more from humans in a similar proportion that computers take more from humans.



 $Prob(HC) > (1 - \alpha_{m}) \alpha_{t}$ or $Prob(CH) > \alpha_{m}(1 - \alpha_{t})$ or $Prob(CC) < \alpha_{m} \alpha_{t}$

If we find that RC/RH>1, then we conclude that computers take more from humans, than humans themselves, in other words, computer trading is more correlated than expected, as computers trade less with other computers than expected or computers trade more with humans than expected.



be estimate:

 $R = RC/RH = \frac{\frac{Vol(HC)}{Vol(CC)}}{\frac{Vol(HH)}{Vol(CH)}}$

At 1-minute, 5-minute and daily frequency

Report ln(R)

If we find that ln(R)>0, then we conclude that computer trading is more correlated than expected



		1-min data			5-min data			Daily data	
	$\ln(R)$	$\ln(R^S)$	$\ln (R^B)$	$\ln(R)$	$\ln (R^S)$	$\ln (R^B)$	$\ln(R)$	$\ln(R^S)$	$\ln (R^B)$
					USD/EUR				
Mean	0.2216***	0.0545***	0.0399***	0.3672***	0.2116***	0.2074***	0.531***	0.4896***	0.4993***
(std. err.)	(0.0031)	(0.0042)	(0.0042)	(0.0036)	(0.0046)	(0.0047)	(0.0118)	(0.0122)	(0.0118)
Percent of obs.>0	0.599	0.527	0.522	0.722	0.627	0.626	0.99	0.975	0.974
No. of non-missing obs.	143539	89960	87597	52174	44366	43609	881	847	855
Total no. of obs.	512640	512640	512640	102528	102528	102528	1068	1068	1068
					JPY/USD				
Mean	0.3106***	0.1535***	0.1513***	0.3923***	0.2119***	0.2111****	0.5846***	0.5755***	0.5398***
(std. err.)	(0.0037)	(0.0052)	(0.0052)	(0.0043)	(0.0054)	(0.0054)	(0.0131)	(0.0143)	(0.0133)
Percent of obs.>0	0.611	0.545	0.545	0.703	0.609	0.61	0.99	0.969	0.971
No. of non-missing obs.	114538	61048	61683	49980	39077	39413	954	939	923
Total no. of obs.	512640	512640	512640	102528	102528	102528	1068	1068	1068
					JPY/EUR				
Mean	0.6984 * **	0.531***	0.5196***	0.6873***	0.584***	0.5837***	0.8129***	0.7736***	0.7376***
(std. err.)	(0.0049)	(0.0074)	(0.0076)	(0.0051)	(0.0066)	(0.0067)	(0.0173)	(0.0171)	(0.0163)
Percent of obs.>0	0.696	0.643	0.636	0.758	0.7	0.698	0.984	0.975	0.965
No. of non-missing obs.	71810	30571	29346	45889	31648	31028	988	952	943
Total no. of obs.	512640	512640	512640	102528	102528	102528	1068	1068	1068



Do algorithmic trades, strategies, tend to be correlated?

Answer: Yes. It seems that computers do not trade with each other as much as random matching would predict.



Relationship between algorithmic trading activity and triangular arbitrage opportunities

Graphical evidence

Percent of seconds with triangular arbitrage profit greater than 1 basis point, in 3-11 time interval



Does algorithmic trading increase or decrease triangular arbitrage opportunities?

- Endogeneity (reverse causality) problem: Triangular arbitrage must clearly also cause AT
 - Granger Causality at high frequency (minute-by-minute) and
 - Heteroskedasticity identification



Structural VAR Estimation

$$AY_t = \Phi(L)Y_t + \Lambda X_{t-1:t-20} + \Psi G_t + \varepsilon_t$$

$$Y_t = (Arb_t, AT_t)$$

- $\Phi(L)$ Lag-polynomial, 20-lags
- $X_{t-1:t-20}$ Controls for past volatility and liquidity (volume)
- G_t deterministic intra-daily patterns and time trend

AT participation causes triangular arbitrage opp.

Tests of VAT Causing Triangular	Arbitrage
Sum of coeffs. on VAT lags	-0.0027***
Sum of coeffs. on VAT lags $ imes \sigma_{VAT}$	-0.0694***
$\chi_1^2 (Sum = 0)$	7.7079
p-value	0.0055
χ^2_{20} (All coeffs. on VAT lags = 0)	91.4066***
p-value	0.0000
Contemp. coeff.	-0.0145***
(std. err.)	(0.0040)
Contemp. coeff.× σ_{VAT}	-0.3731***
No. of obs.	512016
No. of obs. in 1st sub semple	256246
No. of obs. in 2nd sub semple	255770
No. of unique days in 1st sub sample	534
No. of unique days in 2nd sub sample	533



AT liquidity demand causes triangular arbitrage opportunities

Tests of OFCt Causing Triangular	Arbitrage
Sum of coeffs. on OFCt lags	-0.0117***
Sum of coeffs. on $OFCt$ lags $\times \sigma_{OFCt}$	-0.2601***
$\chi_1^2 (Sum = 0)$	142.4661
p-value	0.0000
χ^2_{20} (All coeffs. on $OFCt$ legs = 0)	317.3994***
p-value	8000.0
Contemp. coeff.	-0.0288***
(std. err.)	(0.0055)
Contemp. coeff.× ooFCt	-0.6226***
No. of obs.	511688
No. of obs. in 1st sub sample	255972
No. of obs. in 2nd sub sample	255716
No. of unique days in 1st sub sample	534
No. of unique days in 2nd sub sample	533



AT liquidity supply causes triangular arbitrage opportunities

Tests of VCm Causing Triangular	Arbitrage
Sum of coeffs. on VCm lags	0.0049***
Sum of coeffs. on VCm lage $\times \sigma_{VCm}$	0.0763***
$\chi_1^2 (Sum = 0)$	14.4351
p-value	0.0001
χ^2_{20} (All coeffs. on VCm lags = 0)	46.5108***
p-value	0.0007
Contemp. coeff.	-0.0040***
(std. err.)	(0.0011)
Contemp. coeff. $\times \sigma_{VCm}$	-0.0623***
No. of obs.	512018
No. of obs. in 1st sub sample	256246
No. of obs. in 2nd sub semple	255770
No. of unique days in 1st sub sample	534
No. of unique days in 2nd sub sample	533



High AT correlated actions causes triangular arbitrage opportunities

Tests of $\ln(R)$ Causing Triang	ular Arbitrage
Sum of coeffs. on $\ln(R)$ lags	-0.0437
Sum of coeffs. on $\ln{(R)} \lg \times \sigma_{\ln{(R)}}$	-0.0288
$\chi_1^2 (Sum = 0)$	1.6112
p-value	0.2043
χ^2_{20} (All coeffs. on $\ln(R)$ lags = 0)	24.8891
p-value	0.2057
Contemp. coeff.	-0.8782**
(std. err.)	(0.4413)
Contemp. coeff.× $\sigma_{ln(R)}$	-0.5747**
No. of obs.	147114
No. of obs. in 1st sub sample	8530
No. of obs. in 2nd sub sample	138584
No. of unique days in 1st sub sample	184
No. of unique days in 2nd sub sample	533



triangular arbitrage opportunities causes AT "signed" liquidity demand

Tests of Triangular Arbitrage (Causing $Of(Ct)$
Sum of coeffs. on arb lags	0.0345***
Sum of coeffs. on arb lags $\times \sigma_{arb}$	0.0345***
$\chi_1^2(Sum = 0)$	75.9510
p-value	0.0000
χ^2_{20} (All coeffs. on arb lags = 0)	459.3056***
p-value	0.0008
Contemp. coeff.	1.3037***
(std. err.)	(0.1241)
Contemp. coeff. $\times \sigma_{arb}$	4.4325***
No. of obs.	511688
No. of obs. in 1st sub sample	255972
No. of obs. in 2nd sub sample	255716
No. of unique days in 1st sub sam	ple 534
No. of unique days in 2nd sub san	aple 533

Triangular Arbitrage Causality Tests

- At reduces triangular arbitrage opportunities
- Predominantly AT acts on posted quotes by other traders that enable the profit opportunity
- Increase the speed of price discovery, but increase adverse selection costs of slow traders
- Some evidence that algorithmic traders make prices more efficient by posting quotes that reflect new information quickly



Does algorithmic trading increase or decrease "excess" volatility: autocorrelation of high frequency returns?

Graphical evidence



5-second return autocorrelation





5-second return autocorrelation

JPY/USD 0.5 0.45 0.4 0.35 0.3 0.25 0.2 0.15 0.1 0.05 0 Jan-03 Jan-04 Jan-05 Jan-06 Jan-07

35



5-second return autocorrelation



AT participation causes HF return autocorrelation

	USD/EUR	JPY/USD	JPY/EUR
Tests of VAT	Causing Autocor	relation	
Sum of coeffs. on VAT lags	-0.03904***	-0.02562***	-0.01539***
Sum of coeffs. on VAT lags $ imes \sigma_{VAT}$	-0.93485***	-0.73765***	-0.57837***
$\chi_1^2 (\text{Sum} = 0)$	33.7507	25.2649	14.9665
p-value	0.0000	0.0000	0.0001
χ^2_{20} (All coeffs. on VAT lags = 0)	43.0591***	45.5455***	15.1881***
p-value	0.0000	0.0000	0.0043
Contemp. coeff.	-0.04770***	-0.02377***	-0.03274***
(std. err.)	(0.00703)	(0.00622)	(0.00927)
Contemp. coeff. $\times \sigma_{VAT}$	-1.14222***	-0.68439***	-1.23041***
No. of obs.	102432	102427	102113
No. of obs. in 1st sub sample	51264	51260	51019
No. of obs. in 2nd sub semple	51168	51167	51094
No. of unique days in 1st sub sample	534	534	534
No. of unique days in 2nd sub sample	533	533	533

AT liquidity demand causes HF return autocorrelation

	USD/EUR	: JPY/USD	JPY/EUR	_
Tests of OFCt (Causing Autocor	relation		
Sum of coeffs. on OFCt lags	-0.00481	-0.00154	0.00013	-
Sum of coeffs. on $OFCt$ lags× σ_{OFCt}	-0.12145	-0.04392	0.004664	
χ_1^2 (Sum = 0)	2.3004	0.2637	0.0021	
p-value	0.1293	0.6076	0.9638	
χ^2_{20} (All coeffs. on OFCt lags = 0)	2.8266	1.8544	1.3831	
p-value	0.5872	0.7625	0.8471	_
Contemp. coeff.	0.00434	-0.00631	-0.00124	
(std. err.)	(0.00455)	(0.00463)	(0.004€4)	
Contemp. coeff.× oorCt	0.109584	-0.17996	-0.04448	
No. of obs.	102250	101992	100815	
No. of obs. in 1st sub sample	51095	50871	49935	
No. of obs. in 2nd sub semple	51155	51121	50880	
No. of unique days in 1st sub sample	534	534	534	
No. of unique days in 2nd sub sample	533	533	533	

AT liquidity supply causes HF return autocorrelation

	TIOD (THID)	1011/1100	TTALE (TALES
	USD/EUR	JPY/USD	JPY/EUR
Tests of VCm C	ausing Autocom	relation	
Sum of coeffs. on VCm lags	-0.04692***	-0.03620***	-0.02146***
Sum of coeffs. on VCm lags× σ_{VCm}	-0.73141***	-0.70842^{***}	-0.55656***
$\chi_1^2 (\text{Sum} = 0)$	29.7277	28.8421	17.8840
p-value	0.0000	0.0000	0.0000
χ^2_{20} (All coeffs. on VCm lags = 0)	44.1685***	38.8638***	18.9016***
p-value	0.0000	0.0000	0.0008
Contemp. coeff.	-0.07073***	-0.03907***	-0.02409***
(std. err.)	(0.00902)	(0.00711)	(0.00€22)
Contemp. coeff.× σ_{VCm}	-1.10257***	-0.76458***	-0.62477***
No. of obs.	102432	102427	102113
No. of obs. in 1st sub sample	51264	51260	51019
No. of obs. in 2nd sub sample	51168	51167	51094
No. of unique days in 1st sub sample	534	534	534
No. of unique days in 2nd sub sample	533	533	533



High AT correlated actions causes HF return autocorrelation

	USD/EUR	JPY/USD	JPY/EUR
Tests of ln (R)	Causing Autocor	relation	
Sum of coeffs. on $\ln(R)$ lags	-0.22045^{*}	-0.10045	0.19150*
Sum of coeffs. on $\ln(R)$ lags× σ_{OFCt}	-0.25828	-0.12631	0.249651
$\chi_1^2 (Sum = 0)$	3.6279	0.9306	3.2797
p-value	0.0568	0.3347	0.0701
χ^2_{20} (All coeffs. on $\ln(R)$ lags = 0)	5.6540	1.8232	3.8823
p-value	0.2265	0.7682	0.4222
Contemp. coeff.	0.32317	0.57153	-0.00193
(std. err.)	(0.26693)	(0.41419)	(15.11631)
Contemp. coeff.× σOFCt	0.37863	0.718661	-0.00252
No. of obs.	50166	46115	38653
No. of obs. in 1st sub sample	5860	5258	5236
No. of obs. in 2nd sub sample	44306	40857	33417
No. of unique days in 1st sub sample	255	329	335
No. of unique days in 2nd sub sample	533	533	533

HF return autocorrelation causes AT liquidity supply

	USD/EUR	JPY/USD	JPY/EUR
Tests of Autocor	relation Causing	y VCm	
Sum of coeffs. on ac lags	-0.00917***	-0.01358***	-0.02662***
Sum of coeffs. on ac lage $ imes \sigma_{ac}$	-0.00157***	-0.00235***	-0.0048***
$\chi_1^2 (Sum = 0)$	9.3691	10.1188	20.1782
p-value	0.0022	0.0015	0.0000
χ^2_{20} (All coeffs. on ac lags = 0)	33.8031***	11.0068**	20.2980***
p-value	0.0000	0.0265	0.0004
Contemp. coeff.	0.0015€	0.00268	0.03122***
(etd. +>>.)	(0.00137)	(0.00260)	(0.00571)
Contemp. coeff. $\times \sigma_{\alpha c}$	0.000267	0.000463	0.005635***
No. of obs.	102432	102427	102113
No. of obs. in 1st sub sample	51264	51260	51019
No. of obs. in 2nd sub semple	51168	51167	51094
No. of unique days in 1st sub sample	534	534	534
No. of unique days in 2nd sub sample	533	533	533

Conclusion

- We find evidence of algorithmic trading improving price efficiency:
 - Reduces triangular arbitrage opportunities: mainly by acting on the posted quotes of other traders that enable the profit opportunity
 - Reduces HF return autocorrelation: mainly by providing liquidity
- Caveat 1: Algorithmic trades tend to be correlated, and when this happens we find higher HF return autocorrelation, although the effect is not statistically significant.
- Caveat 2: We do not have truly turbulent times in our sample. We look forward to analyzing data during the crisis.



Future Research

- EBS in 2009 imposed Minimum Quote Life (250 miliseconds) to promote AT's intention to trade
- How did HFT behave during the crisis in the government bond market?



Backup Slides

Theory (continued)

- Foucault, Kadan, and Kandel (2009) model AT as lowering monitoring costs
 - Pareto optimal (lower trading costs, increase trading rate)
 - Ambiguous effect on bid-ask spread (liquidity)
 - When monitoring costs for market-makers ↓ liquidity ↑ informational efficiency ↑
 - When monitoring costs for market-takers \downarrow liquidity \downarrow informational efficiency \downarrow

We have four events: H-make, C-make, H-take, C-take. The probability of each event at time *k* is

 $Prob(C-take) = \alpha_t$, $Prob(H-take) = 1 - \alpha_t$

 $Prob(C-make) = \alpha_m$, $Prob(H-make) = 1 - \alpha_m$

Assuming each event is independent, the probabilities of each trading event are:

 $Prob(HH) = (1 - \alpha_m)(1 - \alpha_t)$ $Prob(HC) = (1 - \alpha_m) \alpha_t$ $Prob(CH) = \alpha_m(1 - \alpha_t)$ $Prob(CC) = \alpha_m \alpha_t$

 We can write the following identities: Prob(CH)×Prob(HC) = Prob(CC)×Prob(HH)
 Prob(HC)/Prob(CC) = Prob(HH)/ Prob(CH)



Taylor Expansion

	1-min data				5-min data			Daily data		
	$\ln(R)$	$\ln (R^S)$	$\ln (R^B)$	$\ln(R)$	$\ln (R^S)$	$\ln (R^B)$	$\ln(R)$	$\ln (R^{S})$	$\ln (R^B)$	
					USD/EUR					
Mean	0.8492***	0.8088***	0.8306***	0.7601***	0.6721***	0.6923***	0.8644***	0.8115***	0.8057***	
(std. err.)	(0.0023)	(0.0028)	(0.0028)	(0.0036)	(0.0042)	(0.0042)	(0.0314)	(0.0297)	(0.0297)	
Percent of obs.>0	0.779	0.736	0.736	0.814	0.758	0.763	1	0.979	0.978	
No. of non-missing obs.	422880	406560	410400	84576	81312	82080	1001	981	978	
Total no. of obs.	512640	512640	512640	102528	102528	102528	1068	1068	1068	
					JPY/USD					
Mean	0.9784***	0.954***	0.9134***	0.8755***	0.8375***	0.7986***	0.6631***	0.6601***	0.6205***	
(std. err.)	(0.0026)	(0.0034)	(0.0034)	(0.0039)	(0.0046)	(0.0046)	(0.0173)	(0.0173)	(0.0165)	
Percent of obs.>0	0.774	0.735	0.727	0.814	0.768	0.763	0.99	0.971	0.972	
No. of non-missing obs.	457920	450720	443040	91584	90144	88608	1000	1000	981	
Total no. of obs.	512640	512640	512640	102528	102528	102528	1068	1068	1068	
	JPY/EUR									
Mean	1.1261***	1.0458***	1.005***	1.1027***	1.0267***	0.9875***	0.8734***	0.855***	0.8456***	
(std. err.)	(0.0033)	(0.0045)	(0.0046)	(0.0045)	(0.0053)	(0.0053)	(0.0181)	(0.0184)	(0.0191)	
Percent of obs.>0	0.772	0.745	0.732	0.827	0.786	0.777	0.985	0.973	0.968	
No. of non-missing obs.	474240	456960	452640	94848	91392	90528	1061	1061	1020	
Total no. of obs.	512640	512640	512640	102528	102528	102528	1068	1068	1068	



Heteroskedasticity identification

 $y = \beta_1 x + \varepsilon$ $x = \beta_2 y + \eta$

• 4 parameters: $\beta_1, \beta_2, \sigma_{\varepsilon}, \sigma_{\eta}$

- 3 moments: Var(y), Var(x), Cov(x, y)
- System is not identified



Heteroskedasticity identification

The woregimes, keep coefficients constant across regimes

 $y_{1} = \beta_{1}x_{1} + \varepsilon_{1}$ $y_{2} = \beta_{1}x_{2} + \varepsilon_{2}$ $x_{1} = \beta_{2}y_{1} + \eta_{1}$ $x_{2} = \beta_{2}y_{2} + \eta_{2}$

• 6 parameters: β_1 , β_2 , $\sigma_{\varepsilon 1}$, $\sigma_{\eta 1}$, $\sigma_{\varepsilon 2}$, $\sigma_{\eta 2}$

6 moments:
Var(y_1), Var(y_2), Var(x_1), Var(x_2), Cov(x_1 , y_1), Cov(x_2 , y_2

System is exactly identified