The Academy of Economic Studies
The Faculty of Finance, Insurance, Banking and Stock Exchange
Doctoral School of Finace and Banking

Dynamics of Exchange Rates Using Inhomogenous Tick-by-tick Data. The Case of the EURRON Currency Pair.

MSc Student: Vlad Mihai Petrescu
Supervisor: Professor Moisă Altăr, PhD

JUNE 2012, Bucharest

Contents

Motivation

- As financial markets continue to deepen and grow and significant advances in technology allow trading venues to conduct transactions with ever smaller latencies, insights from market microstructure will become more important
- The highest sampling frequency is tick-by-tick data; recent literature on models based on asymmetric information theory suggests that time has a role in the dissemination of information among inhomogeneous participants to trading
- Standard econometric techniques applied to high frequency data are sensitive to interpolation techniques
- Implications for institutional investors' trading book and risk management

Brief Literature Survey

- O'Hara (1995) note that "adjustment of prices to information depends on time"
- Engle and Russell (1998) proposes the Autoregressive Conditional Duration (ACD) model to account for duration clustering, standardized durations follow an exponential distribution; Engle (2000) extends the model to the Weibull distribution; Lunde (1999) proposes the Gamma distribution
- Jasiak (1998) propose the Fractionally Integrated ACD (FIACD) model
- Engle (2000) extend the models to the price adjustment process by using an ACD-GARCH framework

- Joint density and log-likelihood for $\{(x_i, z_i), i = 1, \dots, T\}$

$$f\left(x_i, z_i | \breve{x}_{i-1}, \breve{z}_{i-1}; \theta_f\right) = g(x_i | \breve{x}_{i-1}, \breve{z}_{i-1}; \theta_x) q(z_i | x_i, \breve{x}_{i-1}, \breve{z}_{i-1}; \theta_z)$$

$$\mathcal{L}(\theta_x, \theta_z | \mathcal{F}_T^{x,z}) = \sum_{i=1}^{T} [\log g(x_i | \breve{x}_{i-1}, \breve{z}_{i-1}; \theta_x) + \log q(z_i | x_i, \breve{x}_{i-1}, \breve{z}_{i-1}; \theta_z)]$$

- Durations are considered weakly exogenous

- Intraday seasonal component

$$s\left(\theta_s, m(t_i)\right) = \mathrm{e}^{h\left(\theta_s, m(t_i)\right)}$$

$$x_i = \tilde{x}_i s\left(\theta_s, m(t_i)\right)$$

$$\log x_i = \log \tilde{x}_i + h\left(\theta_s, m(t_i)\right)$$

- EACD(p,q) model proposed by Engle and Russell (1998)

$$\Psi_i = \omega + \alpha(L)X_i + \beta(L)\Psi_i$$

- FIACD(p,d,q) model proposed by Jasiak (1998)

$$[1-\beta(L)]\Psi_i = \omega + [1-\beta(L)-[1-\phi(L)](1-L)^d]X_i$$

- ACD-GARCH models proposed by Engle (2000)
  - Variance of the transaction versus variance per second $h_i = x_i\sigma_i^2$
  - ARMA(1,1)-GARCH(1,1) specification for $r_i = \tilde{r}_i x_i^{-1/2}$

$$r_i = \rho r_{i-1} + u_i + \phi u_{i-1}$$

$$u_i = \sigma_i z_i \text{ with } z_i \text{ i.i.d} \sim t_\nu(0,1)$$

$$\sigma_i^2 = \omega + \alpha u_{i-1}^2 + \beta\sigma_{i-1}^2$$

  - Conditional variance with external regressor

$$\sigma_i^2 = \omega + \alpha u_{i-1}^2 + \beta\sigma_{i-1}^2 + \gamma x_i^{-1}$$

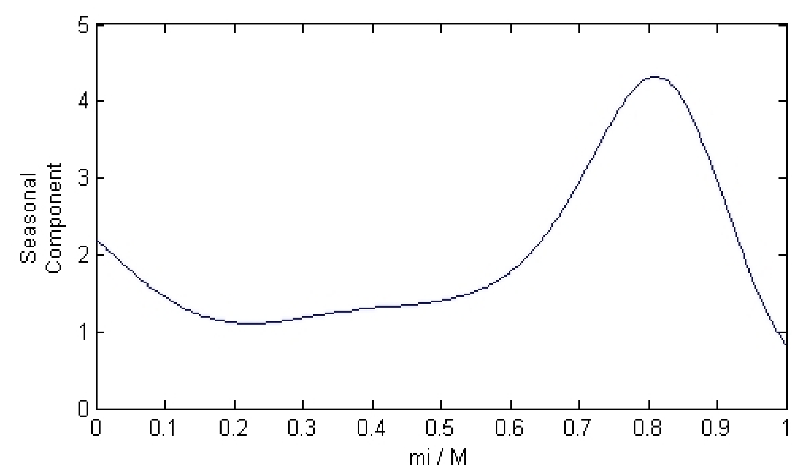
Data and duration seasonality

- EURRON price data provided by Olsen&Associates (timestamp, bid and ask quotes, institution, Olsen filter)
  - 3-Mar-2008 to 31-Mar-2011; 159 trading days after filtering; 360 average trades/day
- Duration seasonality estimated with a Fourier Flexible Form

$$h(\theta_s, m(t_i)) = \mu_1 \frac{m(t_i)}{M} + \mu_2 \left(\frac{m(t_i)}{M}\right)^2 + \sum_{p=1}^{P} \left(\delta_{c,p} \cos\left(2p\pi \frac{m(t_i)}{M}\right) + \delta_{s,p} \sin\left(2p\pi \frac{m(t_i)}{M}\right)\right)$$

Variable	Coefficient		Std. Error	t-Statistic
μ0	3.31	***	0.14	23.55
μ1	5.48	***	0.69	7.91
μ2	-6.46	***	0.67	-9.59
δ cos 01	0.79	***	0.06	14.08
δ sin 01	-0.85	***	0.04	-19.87
δ sin 02	-0.27	***	0.02	-15.12

R-squared	0.0219
Adjusted R-squared	0.0218
Akaike info criterion	3.8056
Schwarz criterion	3.8063
Durbin-Watson stat	1.2639



Long memory in durations - FIACD

I(0) and I(1) tests

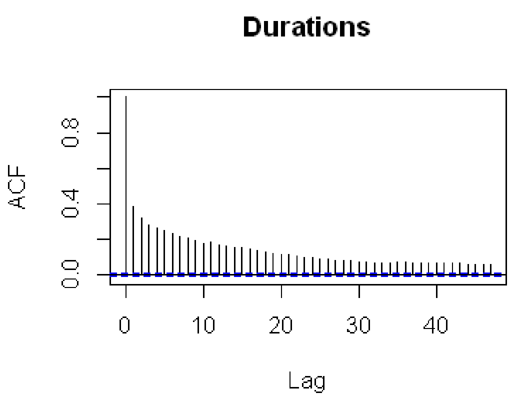
Test	Statistic	Critical Value (0.01)
Durations	-	-
KPSS		
Levels	7.324***	0.739
Trend	0.835***	0.216
ADF test	-28.114***	-4.040
PP test	-206.719***	-4.040
Fractionally Differenced Durations		
KPSS		
Levels	0.657**	0.739
Trend	0.092*	0.216

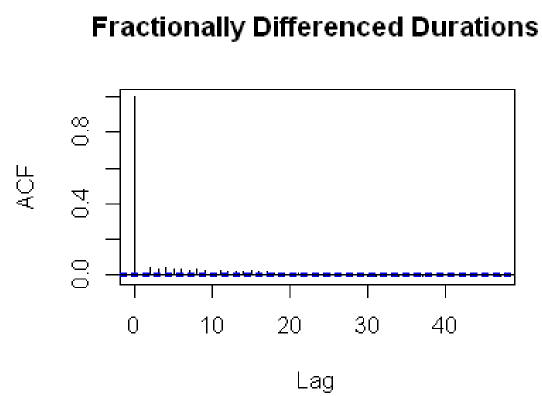
GPH estimator

Test	Statistic	Asymptotic SD	Standard Error Deviation
GPH *d*	0.2384	0.0434	0.0492

Ljung-Box Statistics

	Durations	FD Durations
Q(1)	8314.43***	1.64
Q(5)	26375.34***	313.88***
Q(20)	51365.94***	631.21***





ACD Models (1)

Parameter estimates of ACD models

	EACD(1,1)		WACD(1,1)		GACD(1,1)		
$\hat{\omega}_x$	**0.035**		**0.003**		**0.031**		*Est.*
	(0.023)		(0.040)		(0.028)		*Std. Err.*
	1.519		0.077		1.110		*t-stat.*
$\hat{\alpha}_x$	**0.041**	***	**0.055**	***	**0.041**	***	*Est.*
	(0.003)		(0.004)		(0.003)		*Std. Err.*
	16.178		13.953		13.261		*t-stat.*
$\hat{\beta}_x$	**0.907**	***	**0.880**	***	**0.907**	***	*Est.*
	(0.006)		(0.010)		(0.008)		*Std. Err.*
	139.716		91.492		114.740		*t-stat.*
$\hat{\gamma}_x$			**0.757**	***			*Est.*
			(0.002)				*Std. Err.*
			321.287				*t-stat.*
$\hat{\kappa}_x$					**0.679**	***	*Est.*
					(0.003)		*Std. Err.*
					198.692		*t-stat.*

- The Constrained Maximum Likelihood estimation procedure adopted converges for all three models considered, and the parameter estimates for α and ß are significant and add up to nearly one, however we would expect a sum much closer to one for a large sample if the long memory process had not been accounted for through FIACD

ACD Models (2)

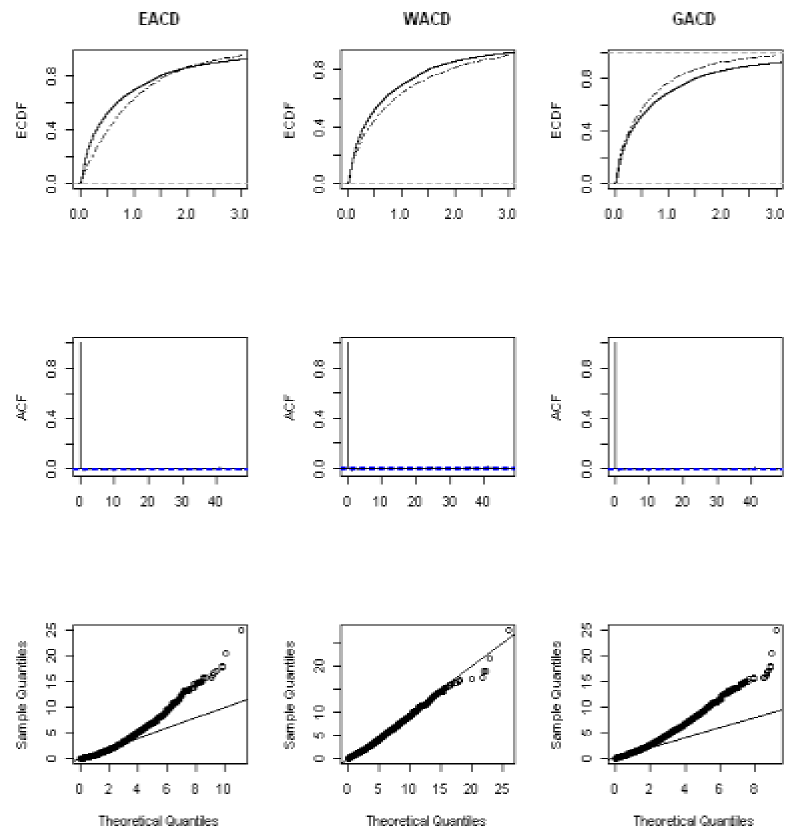
Diagnostic tests

Statistic	EACD(1,1)		WACD(1,1)		GACD(1,1)	
Q(1)	11.8916	***	24.4393	***	11.8916	***
Q(10)	24.2777	***	50.8226	***	24.2777	***
Log-Likelihood	-277985.10		-273313.6		-274663.1	
LR			9343	***	6644	***
KS	0.01426	***	0.0722	***	0.0813	***

- Q-Statistics still show significant serial correlation in standardized durations, however the values have decreased sharply
- Likelihood Ratio tests imply that both Weibull and Gamma distributions are a better fit than the exponential distribution
- Kolmogorov-Smirnov tests suggest that the specification can be improved by more flexible distributional assumptions

ACD Models (3)

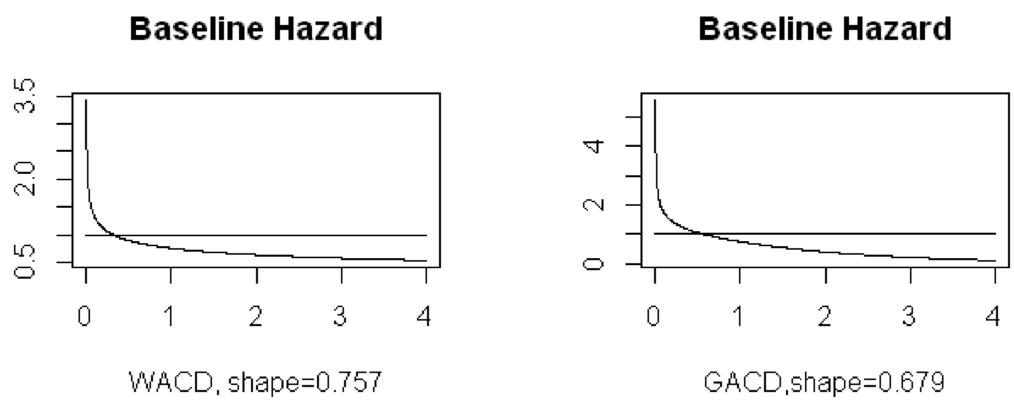
Empirical CDF versus theoretical CDF, ACF plots, QQ plots



- The models considered fit the data reasonably well, except for the tail behavior, however large deviations from theoretical quantiles represent less than 2.5% of the sample

ACD Models (4)

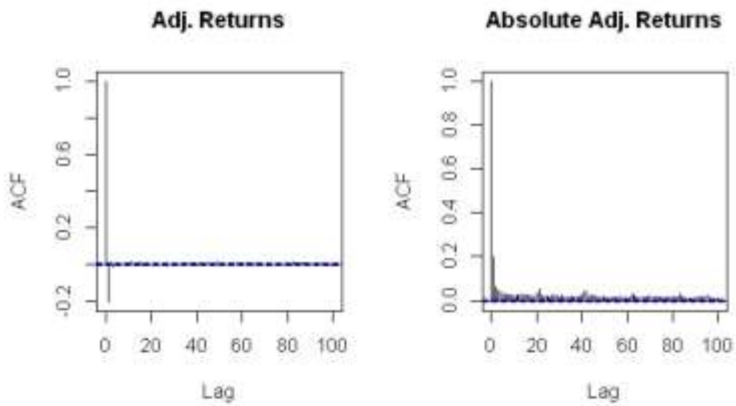
- Baseline hazard function for the Weibull and Standard Gamma distributions versus the constant exponential hazard



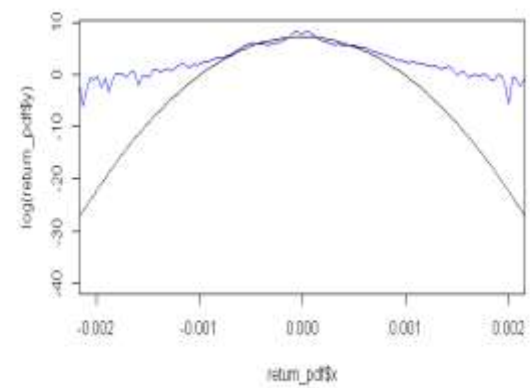
- The estimated parameters of the Standard Weibull and Standard Gamma distributions are robust to starting values for maximum likelihood estimation and imply a strong rejection of the constant hazard, but also of the non-monotonic hazard function – “early failures”

ACD-GARCH Framework (1)

ACF plots for adjusted returns and absolute adjusted returns



Log-PDF of returns against N(0,1)



- ACF of returns suggests MA(1), standatd Student t-distribution can be used to model heavy tails
- Engle (2000) refers to the model as Ultra-High Frequency (UHF)-GARCH

ACD-GARCH Framework (2)

GARCH(1,1) model parameter estimates

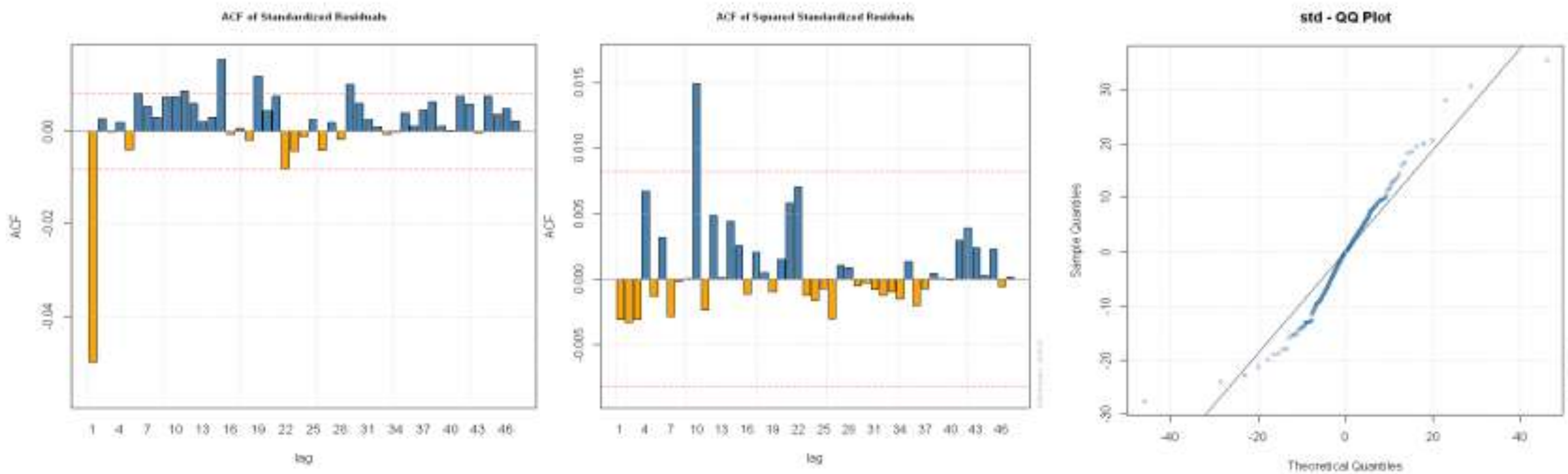
	Estimate	Std. Err.	t Stat.	p-value	Robust Std. Err.	Robust t Stat.	Robust p-value
$\hat{\mu}_r$	0.00007*	0.00003	2.09	0.036	0.00004	1.93	0.053
MA(1)	-0.20428***	0.00378	-54.04	0.000	0.00543	-37.62	0.000
$\hat{\omega}_r$	0.00034***	0.00001	28.21	0.000	0.00002	17.02	0.000
$\hat{\alpha}_r$	0.92992***	0.03018	30.81	0.000	0.03813	24.39	0.000
$\hat{\beta}_r$	0.06908***	0.01053	6.56	0.000	0.02424	2.85	0.004
$\hat{\nu}_r$	2.31107***	0.01144	202.09	0.000	0.00952	242.88	0.000

- In the mean equation, we observe the significant "bid-ask bounce" effect
- The mean equation constant is not significant at 1%, as in many studies such as Engle (2000), Liu (2010), etc.
- We observe the "IGARCH effect" in the conditional duration equation
- The shape parameter estimate confirms the pronounced heavy tail distribution of returns, also suggests that inference should be made considering robust statistics; improvements can be made by using a mixture of distributions
- The following is true:

$$E\left(r_i^2 \middle| \mathcal{F}_{i-1}\right) = E\left(h_i \middle| \mathcal{F}_{i-1}\right) = E\left(x_i \sigma_i^2 \middle| \mathcal{F}_{i-1}\right) = \Psi_i \sigma_i^2$$

ACD-GARCH Framework (3)

GARCH(1,1) model diagnostic plots



- ACF of residuals and squared residuals show that the model is potentially misspecified (as confirmed by portmanteau tests), however the correlations are relatively low
- The Sudent t-distribution manages to account relatively well for the heavy tail behavior

ACD-GARCH Framework (4)

IGARCH(1,1) model parameter estimates

	Estimate	Std. Err.	t Stat.	p-value	Robust Std. Err.	Robust t Stat.	Robust p-value
$\hat{\mu}_r$	0.00007**	0.00003	2.09	0.036	0.00003	2.23	0.026
MA(1)	-0.20429***	0.00384	-53.24	0.000	0.00421	-48.58	0.000
$\hat{\omega}_r$	0.00034***	0.00001	29.26	0.000	0.00001	23.46	0.000
$\hat{\alpha}_r$	0.93089***	0.00777	119.87	0.000	0.01049	88.72	0.000
$\hat{\beta}_r = 1 - \hat{\alpha}_r$	0.06911***	-	-	-	-	-	-
$\hat{v}_r$	2.31089***	0.00955	241.97	0.000	0.01029	224.70	0.000

- The results of the IGARCH(1,1) specification have a similar interpretation to the ones of the standard GARCH(1,1) model
- IGARCH effect becomes an important issue

ACD-GARCH Framework (5)

- Under IGARCH, the conditional variance is analogous to a random walk process with drift
- Nelson(1990) however points out that if $E(\log(\beta + \alpha z_i^2)) < 0$ then the unconditional variance process is strictly stationary and ergodic, but not weakly stationary; the estimated parameters of the IGARCH(1,1) model satisfy this condition (the mean is –0.299) and $\omega > 0$
- Under these conditions, the unconditional variance is finite almost surely though increasing
- A shock to the square of the noise process is persistent in probability in the unconditional and conditional variance, but not almost surely

ACD-GARCH Framework (6)

IGARCH(1,1) model with external regressor parameter estimates

$$\sigma_i^2 = \omega + \alpha u_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma x_i^{-1}$$

	Estimate	Std. Err.	t Stat.	p-value	Robust Std. Err.	Robust t Stat.	Robust p-value
$\hat{\mu}_r$	0.00013***	0.00003	5.14640	0.00000	0.00003	5.27540	0.00000
MA(1)	-0.16454***	0.00315	-52.29840	0.00000	0.00372	-44.29360	0.00000
$\hat{\omega}_r$	0.00008***	0.00001	15.72660	0.00000	0.00001	14.64090	0.00000
$\hat{\alpha}_r$	0.93405***	0.00215	434.84404	0.00000	0.00290	344.59210	0.00000
$\hat{\beta}_r = 1 - \hat{\alpha}_r$	0.06596***	-	-	-	-	-	-
$\hat{\gamma}_r$	0.01036***	0.00043	24.12380	0.00000	0.00047	22.04890	0.00000
$\hat{\nu}_r$	2.14554***	0.00576	372.33460	0.00000	0.00613	349.81730	0.00000

- Current durations are informative; the reciprocal of the actual durations process enteres the conditional volatility equation
- The positive estimate of γ is consistent with the model of inhomogeneous agents of Easley and O'Hara (1992) – "no trade means no news"
- Duration seasonality can be interpreted as also accounting for an intraday pattern of volatility

ACD-GARCH Framework (6)

Diagnostic tests

GARCH(1,1)	IGARCH(1.1)	IGARCH(1.1) with ext. reg
LogLikelihood : 157166.8	LogLikelihood : 157167.4	LogLikelihood : 165518.6
Information Criteria	Information Criteria	Information Criteria
------------------------------------	------------------------------------	------------------------------------
Akaike -5.5129	Akaike -5.5129	Akaike -5.8058
Bayes -5.5119	Bayes -5.5119	Bayes -5.8047
Shibata -5.5129	Shibata -5.5129	Shibata -5.8058
Hannan-Quinn -5.5126	Hannan-Quinn -5.5126	Hannan-Quinn -5.8055
		LR test 16702.4***

- Including external regressors improves the specification considering information criterion and the likelihod ratio test

Conclusion

- We find evidence that the joint process of durations and prices for the EURRON currency pair has a complex structure which has implications for forecasting and risk management applications at the intraday level
- Considering durations weakly exogenous, the nonparametric "duration clustering" property is explained by long memory and decreasing conditional intensity functions of standardized durations
- Durations between transactions exhibit an intraday seasonal pattern
- Duration-adjusted returns exhibit the familiar "heavy tails" and "duration clustering" properties, as well as microstructure-specific effects such as "bid-ask bounce" and persistence of shocks to the noise process of the IGARCH(1,1) specification in the forecast distribution of long term and conditional variance
- We find evidence that the data is consistent with the model of inhomogeneous agents

Further Improvements

- The models considered are sensitive to structural change, which is a promising area for further research
- Duration seasonality can be improved by modeling interday seasonality
- More flexible distributions for standardized durations
- Simultaneous estimation of more sophisticated models where durations are endogeneous; Granger causality tests

Andresen, T. G. and T. Bollerslev (1998), "Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies," Journal of Finance, American Finance Association, vol. 53(1), pages 219-265, 02

Banerjee, A., J. J. Dolado, J. W. Galbraith, and D. F. Hendry (1993): Cointegration, Error Correction, and the Econometric Analysis of Non-Stationary Data, Oxford University Press, Oxford.

Bauwens, L. and D. Veredas (2004), "The Stochastic Conditional Duration Model: A Latent Factor Model for the Analysis of Financial Durations," Journal of Econometrics, 119, 381-412.

Christian, Y. R. and Rosenbaum, M. (2011), "A New Approach for the Dynamics of Ultra-High-Frequency Data: The Model with Uncertainty Zones", Journal of Financial Econometrics, Vol. 9, 344-366

D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin (1992): Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. Journal of Econometrics 54, 159–178.

Dacorogna, M. M., R. Gencay, U. Muller, R. Olsen and O. Pictet (2001), "An Introduction to High-Frequency Finance", San Diego, CA: Academic Press, 2001

Easley, D., O'Hara, M. (1992) "Time and the process of security price adjustment", Journal of Finance 47, 905-927.

Engle, R. F. (2000), "The Econometrics of Ultra-High Frequency Data", Econometrica, Econometric Society, vol. 68(1), pages 1-22, January.

Engle, R. F. and T. Bollerslev (1986), "Modeling the persistence of conditional variances", Econometric Reviews 5, 1-50

Engle, R. F. and J.R. Russell (2009), "Analysis of High Frequency Financial Data", Handbook of Financial Econometrics, Vol 1.

Engle, R. F. and J.R. Russell (1998), "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," Econometrica, Econometric Society, vol. 66(5), pages 1127-1162, September.

Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long memory time series models, Journal of Time Series Analysis 4(4), 221–238.

Ghysels, E. , C. Gourieroux and J. Jasiak, (2004), "Stochastic volatility duration models," Journal of Econometrics, Elsevier, vol. 119(2), pages 413-433, April.

Ghysels, E. and J. Jasiak (1998): "GARCH for irregularly spaced financial data: the ACD-GARCH model", Studies in Nonlinear Economics and Econometrics, Vol. 2 No. 4, 133-149

Grammig, J and K. Maurer (2000), "Non-monotonic hazard functions and the autoregressive conditional duration model," Econometrics Journal, Royal Economic Society, vol. 3(1), pages 16-38

Grammig, J. & M. Wellner (2002), "Modeling the interdependence of volatility and inter-transaction duration processes," Journal of Econometrics, Elsevier, vol. 106(2), pages 369-400

Jasiak, J. (1998), "Persistence in Intertrade Durations," Finance, 19, 166-195.

Jensen, A. T. and T. Lange (2010): "On Convergence of the QMLE for Misspecified GARCH Models," Journal of Time Series Econometrics, Berkeley Electronic Press, vol. 2(1), pages 3.

Li, G and W. K. Li, (2008) "Least absolute deviation estimation for fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity," Biometrika, Oxford University Press for Biometrika Trust, vol. 95(2), pages 399-414.

Liu, C. and J. M. Maheu (2010): "Intraday Dynamics of Volatility and Duration: Evidence from the Chinese Stock Market," Working Papers tecipa-401, University of Toronto, Department of Economics.

Lunde, A. (1999), "A Generalized Gamma Autoregressive Conditional Duration Model," discussion paper, Aalborg University.

Mariano, R. S. and Tse, Y. K. (2008), "Econometric Forecasting and High-Frequency Data Analysis", Vol. 13, World Scientific

Meitz, M. & T. Terasvirta (2006), "Evaluating Models of Autoregressive Conditional Duration," Journal of Business & Economic Statistics, American Statistical Association, vol. 24, pages 104-124, January.

Morna, C. (2002), "IGARCH effects: an interpretation", Applied Economics Letters, vol. 9, p. 745-748

Nelson, D. (1990), "Stationarity and persistence in the GARCH(1,1) models", Econometric Theory, 6, 318-334

O'Hara, M. (1995), "Market Microstructure Theory", Basil Blackwell, Oxford, England

Pacurar, M. (2006), "Autoregressive Conditional Duration Models In Finance: A Survey Of The Theoretical And Empirical Literature", Journal of Economic Surveys, Wiley Blackwell, vol. 22

Racicot, F. E., R. T. and A. Coen (2006), "Forecasting Irregularly Spaced UHF Financial Data: Realized Volatility vs UHF-GARCH Models," International Advances in Economic Research, Springer, vol. 14(1), pages 112-124

Rodrigues, M. M. P. and A. M. Robert Taylor (2009), "The Flexible Fourier Form and Local GLS De-trended Unit Root Tests," , Working Papers w200919, Banco de Portugal, Economics and Research Department.

Selkuk, F. and Gencay, R. (2006), "Intraday dynamics of stock market returns and volatility", Physica A: Statistical Mechanics and its Applications, Vol. 367, p. 375-387

Tsay, R. S. (2010), "High-Frequency Data Analysis and Market Microstructure", Analysis of Financial Time Series, Third Edition, Third Edition, John Wiley & Sons, Inc., Hoboken, NJ, USA.

Zumbach, G. (2004), "Volatility processes and volatility forecast with long memory", Quantitative Finance, Taylor and Francis Journals, vol. 4(1), pages 70-86.