Advanced Quantitative Methods:

High Frequency Financial Econometrics - A Selection

Ingmar Nolte

Room: C48

I.Nolte@lancaster.ac.uk

Office Hours: whenever you like

Webpage:

http://www.lancaster.ac.uk/staff/nolte/

Password: 42GARCH

- data is available on higher frequencies than daily
- Engle (2000): “ultra high frequency” all single events are recorded
- more information
- individuals become more and more important

$\Rightarrow$ field of high frequency finance and market microstructure analysis

Why is it important?

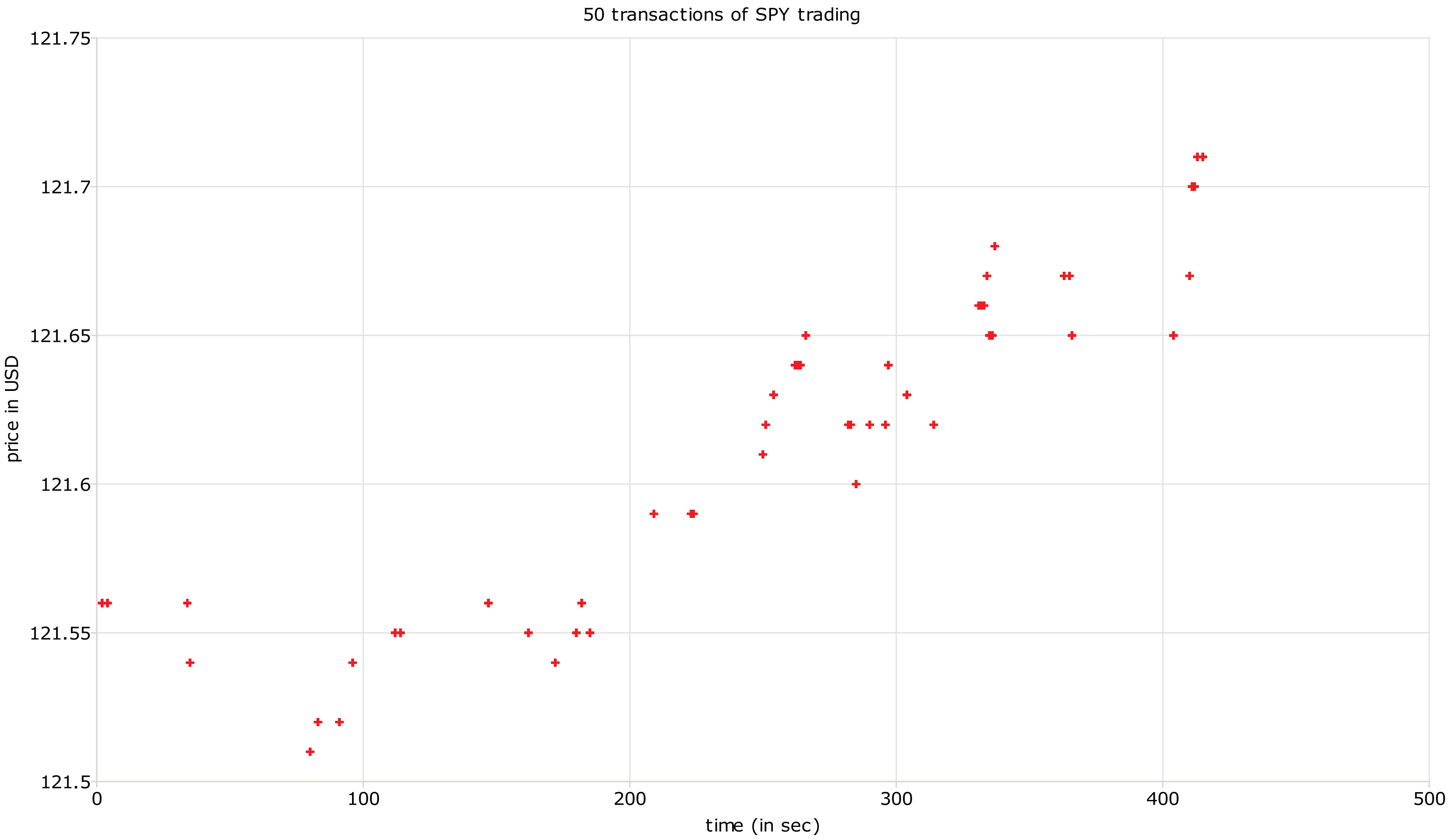
- Older “finance” models can be refined (volatility, asset pricing, ...)
- intraday risk management, intraday trading models,
- analysis of trading behavior,
- design of markets and exchanges (e.g. tick size change NYSE 26^{th} January 2001)

Aim: Understanding the mechanisms at the microlevel

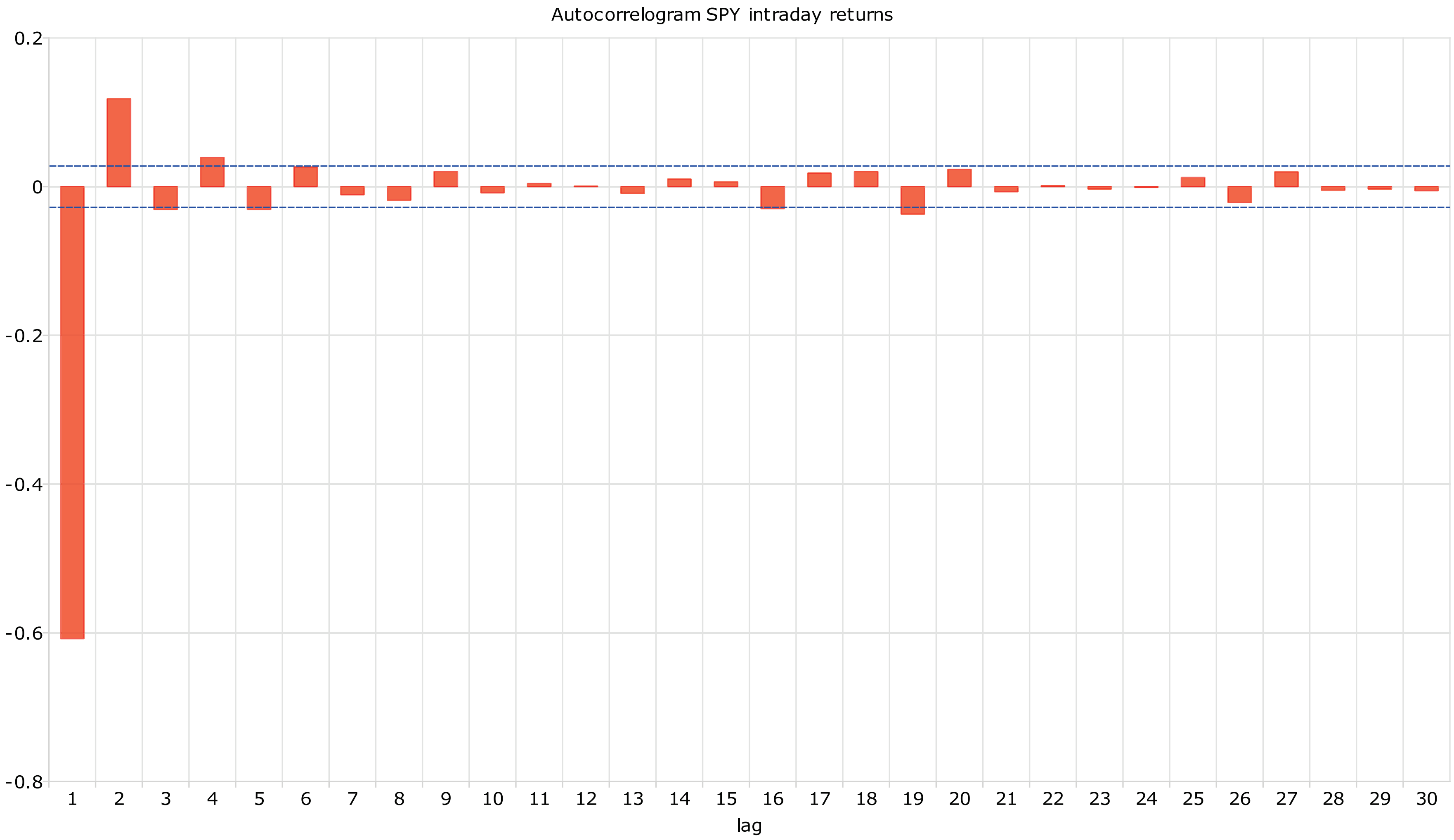
Requirement: Good data quality and appropriate models

2 days of SPY trading

50 observations of SPY trading



Autocorrelogram: SPY



- $t \in \mathbb{R}^+$ continuous physical calendar time.
- $\{t_i\}_{i=1,2,3,\ldots}$ random arrival times of events with $0 \leq t_i \leq t_{i+1}$ on $\mathbb{R}^+$
- $\{t_i\}_{i=1,2,3,\ldots}$ is called point process on $\mathbb{R}^+$
- point process $\{t_i\}_{i=1,2,3,\ldots}$ is simple if $t_i < t_{i+1}$ almost surely

- right continuous counting function for $\{t_i\}_{i=1,2,3,\ldots}$ is given by $N(t) = N([0,t]) = \sum_i \mathbf{1}_{\{t_i \le t\}}$
- left continuous counting function is given by $\breve{N}(t) = N([0,t)) = \sum_i \mathbf{1}_{\{t_i < t\}}$

We assume:

$\{t_i\}_{i=1,2,3,\ldots}$ is integrable, i.e. for all t: $\mathrm{E}[N(t)] < \infty$.

Events?

- $\{W_i\}_{i=1,2,3,\dots}$, with $W_i \in \{1,\dots,K\}$, determines whether the t_i arrival time is induced by the event of type $k \in \{1,\dots,K\}$
- $\{t_i, W_i\}_{i=1,2,3,\dots}$ is then called a K variate marked point process
- representation of K univariate point processes t_i^k, with $N^k(t) = \sum_i \mathbf{1}_{\{t_i \leq t\}} \mathbf{1}_{\{W_i = k\}}$.
- an event can be related to the vector of explanatory variables Z_i
- natural filtration of $N(t)$: $\mathfrak{F}_t^N \equiv \sigma(N(s)|0 \leq s \leq t)$
- broader filtration (covariate process $\{Z_i\}_{i\in\{1,2,\dots\}}$): $\mathfrak{F}_t$ with $\mathfrak{F}_t^N \subseteq \mathfrak{F}_t$.

Consider: univariate point processes $\{t_i\}_{i=1,2,3,...}$

Intensities, Compensators and Durations

- $N(t)$ is an $\mathfrak{F}_t$-submartingale, because $\mathrm{E}[|N(t)|] < \infty$ and since for all $s, t \geq 0$ with $s \leq t$ the condition $\mathrm{E}[N(t)|\mathfrak{F}_s] \geq \mathrm{E}[N(s)]$

Doob-Meyer decomposition of $N(t)$:

$$N(t) = M(t) + \tilde{\Lambda}(t). \tag{1}$$

with

- $\tilde{\Lambda}(t)$ a $\mathfrak{F}_t$-predictable increasing compensator process,
- $M(t)$ a zero mean martingale process,

The $\mathfrak{F}_t$**-intensity process** of $N(t)$ can then be defined as that positive $\mathfrak{F}_t$-predictable process $\lambda(t)$ fulfilling

$$\tilde{\Lambda}(t) = \int_{(0,t]} \lambda(u) du. \tag{2}$$

Theorem of Aalen's (1976):

With $\{\lambda(t)\}$ is bounded from above by an integrable random variable, we get denoting $\lambda(t^+) \equiv \lim_{\Delta \to 0} \lambda(t + \Delta)$ that

$$\lambda(t^+) = \lim_{\Delta \to 0} \frac{1}{\Delta} \mathrm{E}[N(t+\Delta) - N(t) | \mathfrak{F}_t]. \tag{3}$$

Moreover, assuming that $\lambda(t)$ is continuous we naturally obtain that $\lambda(t^+) = \lambda(t)$.

$x_i = t_i - t_{i-1}$: the **inter event duration**, i.e. the time between t_{i-1} and t_i

hazard function:

$$h(x_i) = \lim_{\Delta \to 0} \frac{1}{\Delta} \Pr[x_i \leq X_i < x_i + \Delta | X_i \geq x_i] = \frac{f(x_i)}{1 - F(x_i)}, \tag{4}$$

where x_i is considered as a realization of the random variable X_i, which has the density function $f(x_i)$ and the distribution function $F(x_i)$.

$$S(x_i) \equiv 1 - F(x_i) = \exp\left(-\int_0^{x_i} h(s) ds\right), \tag{5}$$

denotes the survivor function.

Example: Homogeneous Poisson Process

Setting $\lambda(t) \equiv \lambda \in \mathbb{R}^+$ defines the homogenous Poisson process: $\Rightarrow$

$$\tilde{\Lambda}(t) = \int_0^t \lambda(u)du = t\lambda. \tag{6}$$

- the number of events in a certain time interval $(a, b]$, say, $N((a, b]) = N(b) - N(a)$ is Poisson distributed with $\Pr[N((a, b]) = k] = \exp(-\lambda(b-a))\lambda(b-a)^k/k!$
- the inter-event durations $x_i = t_i - t_{i-1}$ are independently exponentially distributed with parameter λ.

Three equivalent representations of the same point process:

- intensity
- duration
- counts

Theorem 0.1 (Meyer (1971), [Random Time Change)**:**]

If the counting process $N(t)$ representing the point process $\{t_i\}_{i\in\{1,2,\dots\}}$ has a continuous compensator process $\tilde{\Lambda}(t)$, with $\tilde{\Lambda}(\infty) = \infty$, then the point process formed from the times $\{\tilde{\Lambda}(t_i)\}_{i\in\{1,2,\dots\}}$ is a homogenous Poisson process with a unit Poisson rate, i.e. $\lambda = 1$ ◇

Construction of $\tilde{N}(t)$ representing the transformed homogenous Poisson process unit Poisson rate:

Choose the (unique) $\mathfrak{F}_t$-stopping times $\tau(t)$ through $\int_0^{\tau(t)} \lambda(u)du = t$ and setting $\tilde{N}(t) = N(\tau(t))$.

Most important implication:

The integrated intensities, $\Lambda(\cdot,\cdot)$ say, which are the increments of $\{\tilde{\Lambda}(t_i)\}_{i\in\{1,2,\ldots\}}$, given by

$$\Lambda(t_{i-1},t_i) \equiv \tilde{\Lambda}(t_i) - \tilde{\Lambda}(t_{i-1}) = \int\limits_{t_{i-1}}^{t_i} \lambda(u)du \tag{7}$$

are independently unit exponentially distributed.

Autoregressive Conditional Duration (ACD) Models

Engle & Russell (1998) and Engle (2000)

Consider:

$x_i \equiv \frac{t_i - t_{i-1}}{s(t_i)}$, where $s(\cdot)$ denotes a seasonality function to capture time of day effects

Aim:

Conditional density of $f(x_i|\mathfrak{F}_{t-1})$ which should have a positive support.

Decompose x_i multiplicatively

$$x_i = \varphi_i \varepsilon_i.$$

- conditional mean function φ_i, capturing all the conditioning information
- i.i.d error term ε_i with unit mean

Playground 1: density of ε_i

A standard choice is the unit exponential density

More flexible specifications: Weibull, Burr, generalized Gamma or generalized F density

Playground 2: dynamics of φ_i

Linear Models:

ACD(p,q):

$$(1-\beta_p(L))\varphi_i = \omega + \alpha_q(L)x_i, \tag{8}$$

Explanatory variables can be included e.g. statically in the following form

$$(1-\beta_p(L))\left(\varphi_i - \zeta Z_i\right) = \omega + \alpha_q(L)x_i,$$

For $p = q = 1$ the ACD model is covariance stationary if

$$(\alpha_1 + \beta_1)^2 - \alpha_1^2 \sigma^2 < 1,$$

where $\sigma^2 = \mathrm{V}[\varepsilon_i] < \infty$. The moments are the given by

$$\begin{aligned}
\mathrm{E}(x_i) &= \mu_x = \frac{\omega}{1-\alpha-\beta} \\
\mathrm{V}(x_i) &= \mu_x^2 \sigma^2 \frac{1-\beta^2-2\alpha\beta}{1-(\alpha+\beta)^2-\alpha^2\sigma^2} \\
\rho_1 &= \frac{\alpha(1-\beta^2-\alpha\beta)}{1-\beta^2-2\alpha\beta} \\
\rho_k &= (\alpha+\beta)\rho_{k-1}, \text{ for } k \geq 2
\end{aligned}$$

Fractionally Integrated ACD **FIACD(p,q):**

A stationary process x_i, is called a stationary process with long memory if there is a $d \in (0.5, 1)$ and a constant C such that

$$\lim_{k \to \infty} \frac{\rho_k}{C k^{2d-1}} = 1. \tag{9}$$

This causes the divergence of the series of absolute autocorrelations, i.e. $\sum_{k=-\infty}^{\infty} |\rho_k| = \infty$, where ρ_k denotes the k^{th} autocorrelation coefficient of x_i.

Rewriting (8) as an ARMA(max(p,q),p) model, yields

$$(1-\alpha_q(L)-\beta_p(L))x_i = \omega + (1-\beta_p(L))v_i,$$

where $v_i = x_i - \varphi_i$. The fractionally integrated model is then obtained by

$$(1-\alpha_q(L)-\beta_p(L))(1-L)^d x_i = \omega + (1-\beta_p(L))v_i, \tag{10}$$

where $(1-L)^d$, $0 < d < 1$ is the fractional differencing operator given by

$$(1-L)^d = \sum_{k=0}^{\infty} \varpi_k L^k,$$

with

$$\varpi_k = \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} = \prod_{0<j\leq k} \frac{j-1-d}{j}, \quad k = 0, 1, 2, \ldots$$

and $\Gamma(\cdot)$ denotes the gamma function given by

$$\Gamma(x) \equiv \begin{cases} \int\limits_0^\infty t^{x-1}\exp(-t)dt & \text{if} \quad x > 0, \\ \infty & \text{if} \quad x = 0, \\ x^{-1}\Gamma(1+x) & \text{if} \quad x < 0. \end{cases}$$

Rewriting (10) the FIACD(p, d, q) model is given by

$$\begin{aligned}(1-\beta_p(L))\varphi_i &= \omega + \left[1 - \beta_p(L) - [1-\alpha_q(L) - \beta_p(L)](1-L)^d\right] x_i \\ &\equiv \omega + \gamma_\infty(L) x_i,\end{aligned}$$

Non-Linear ACD Models:

Example:

$$
\begin{aligned}
(\text{LACD}_1): &\quad (1-\beta_p(L))\ln(\varphi_i) = \omega + \alpha_q(L)\ln(\varepsilon_i),\\
(\text{LACD}_2): &\quad (1-\beta_p(L))\ln(\varphi_i) = \omega + \alpha_q(L)\varepsilon_i,\\
(\text{LACD}_3): &\quad (1-\beta_p(L))\ln(\varphi_i) = \omega + \alpha_q(L)\ln(x_i).
\end{aligned}
$$

In general:

$$B(\varphi_i) = A(\varepsilon_i)B(\varphi_{i-1}) + C(\varepsilon_i),$$

where ε_i is a sequence of i.i.d. variables fulfilling certain regularity conditions, $B(\varphi_i)$ is an $\mathbb{R}$-valued process, $B(\cdot)$ a Borel-function and $A(\cdot)$ and $C(\cdot)$ are polynomial functions.

Dynamic Intensity Models

Autoregressive Conditional Intensity (ACI) Models

Russell (1999)

The conditional intensity function $\lambda(t|\mathfrak{F}_t)$:

$$\lambda(t|\mathfrak{F}_t) = \varphi(t)\lambda_0(t)s(t). \tag{11}$$

- $\varphi(t)$ to capture the dynamic structure
- baseline intensity component $\lambda_0(t)$ independent of $\mathfrak{F}_t$
- seasonality component $s(t)$ to capture diurnally seasonality effects

Why?

Easy incorporation of time varying covariates $Z_{t_i^0}$.

They are allowed to change at any time $t_i^0 \in \mathbb{R}^+$.

The dynamic component is specified as

$$\varphi(t) = \exp\left(\tilde{\varphi}_{\breve{N}(t)+1} + Z'_{\breve{N}^0(t)}\tilde{\gamma}\right), \tag{12}$$

with

$$\tilde{\varphi}_i = \omega + \sum_{j=1}^{p} \alpha_j \varepsilon_{i-j} + \sum_{j=1}^{q} \beta_j \tilde{\varphi}_{i-j}, \tag{13}$$

where the innovation term can be either specified as

$$\varepsilon_i = 1 - \Lambda(t_{i-1}, t_i) \tag{14}$$

or

$$\varepsilon_i = -0.5772 - \ln \Lambda(t_{i-1}, t_i). \tag{15}$$

Please note that $\varphi(t)$ only changes between t_{i-1} and t_i because of the time varying covariates. Thus, abstracting from time varying covariates $\lambda(t)$ changes between t_{i-1} and t_i only deterministically due to $\lambda_0(t)$ and $s(t)$.

Baseline intensity $\lambda_0(t) \leftrightsquigarrow$ error term density in the ACD model.

ACD model specification $x_i = \varphi_i \varepsilon_i$ can be rewritten in terms of intensities as

$$\lambda(t|\mathfrak{F}_t) = \lambda_\varepsilon \left(\frac{x(t)}{\varphi_{\breve{N}(t)+1}} \right) \frac{1}{\varphi_{\breve{N}(t)+1}}, \qquad (16)$$

where λ_ε denotes the intensity function of the ACD error term and $x(t) = t - t_{\breve{N}(t)}$ the backward recurrence time, i.e. the time since the last event $t_{\breve{N}(t)}$.

$\lambda_0(t) = const. \leftrightsquigarrow$ exponentially distributed ACD error term.

Weibull parametrization $\lambda_0(t) = \exp(\omega) x(t)^{a-1}$, with $a > 0$: monotone baseline intensity shape

Burr specification $\lambda_0(t) = \exp(\omega) x \frac{(t)^{a-1}}{1+\eta x(t)^a}$, with $a > 0$ and $\eta \geq 0$: non-monotone baseline intensity shape.

Hawkes Models

A general class of Hawkes processes is given by

$$\lambda(t) = \phi(\mu(t) + \sum_{t_i<t} w(t - t_i)), \tag{17}$$

where ϕ denotes a nonlinear function, $\mu(t)$ a deterministic function in t and $w(\cdot)$ a weighting function.

Hawkes (1971) suggested to parameterize $w(\cdot)$ as

$$w(t) = \sum_{j=1}^{p} \alpha_j \exp(-\beta_j t), \tag{18}$$

where $\alpha_j, \beta_j \geq 0$ for $j = 1, \ldots, p$. The α_j are scale parameters, whereas the β_j determine the decay and therefore the influence of past time.

For $p > 1$, the intensity is a superposition of differently parameterized exponentially decaying weighted sums of the backward recurrence times to **all** previous points.

The stationarity condition for this type of Hawkes process is given by $0 < \int_0^\infty w(s)ds < 1$, which implies $\sum_{j=1}^{p} \frac{\alpha_j}{\beta_j} < 1$.

An alterative specification of the weighting functions which induces a hyperbolic, which is slower than an exponential decay is given by

$$w(t) = \frac{H}{(t+\kappa)^b}, \tag{19}$$

with parameters H, κ and $b > 1$.

Maximum Likelihood Estimation

The likelihood function can be based solely on $\lambda(t)$ (Karr (1991)). The log likelihood for a dataset where the physical time $t \in (0, T]$ is given by

$$\ln \mathcal{L}(\theta) = \int_0^T (1 - \lambda(s))ds + \int_{(0,T]} \ln \lambda(s) dN(s) \quad (20)$$

$$= \sum_{i=1}^{n} -\Lambda(t_{i-1}, t_i) + \ln(\lambda(t_i)) + T, \quad (21)$$

where n denotes the number of events in $(0, t]$ and thus $\{t_i\}_{i=1,\dots,n}$.

Specification of the seasonality function:

- linear splines
- cubic splines
- non-parametrically (Nardaraya-Watson)
- flexible Fourier specification

Goodness-of-fit

- assess the dynamic and the distributional properties
- employ standard tools to the generalized residuals which should be iid Exp(1) distributed
- standard tools to check the dynamics specification: (acf, pacf, LB-tests, BDS-tests, conditional moment tests)
- standard tools to check the distributional assumptions: (QQ-plots, density forecast tests, D-test)

Conditional-Moment-Test

Newey (1985) and Tauchen (1985)

One basically tests on the property of a martingale difference sequence (MDS) or an asymptotically equivalent form in models with multiplicative error terms, like the ACD models. Thus:

$$\mathrm{E}[\varrho_t|\mathfrak{F}_{t-1}] = 0, \tag{22}$$

where $\mathfrak{F}_{t-1}$ denotes the information filtration and $\varrho_t \equiv \varrho_t(Y_t, Z_t, \theta_0)$ is a restriction function[a], that depends on the time series under consideration, Y_t, some further explanatory variables, Z_t, and the true k-dimensional parameter vector, θ_0.

The null hypothesis is equation (22) and the idea to construct a test is:

$$\mathrm{E}[B(\mathfrak{F}_{t-1})\varrho_t] = 0 \text{ for all Borel-functions } B(\cdot) \text{ implies that } \mathrm{E}[\varrho_t|\mathfrak{F}_{t-1}] = 0.$$

Hence, the CM-test statistic is based on the empirical counterparts of these unconditional moments, namely on a

[a]In the ACD model context $\varrho_t = Y_t - \varphi_t^Y$ (MDS) or $\varrho_t = \frac{Y_t}{\varphi_t^Y} - 1$.

vector of $\hat{m}_j$ with

$$\hat{m}_j \equiv \frac{1}{T}\sum_{t=1}^{T} \hat{\varrho}_t B_j(\mathfrak{F}_{t-1})$$

where $B_j(\cdot) \in \mathbb{R}^T$ is an element of a **finite** set (J elements) of arbitrary Borel-functions $\{B_j(\cdot)|j = 1, \dots, J\}$.

Denote $\hat{M}$ as the $T \times J$ dimensional matrix consisting of J $\hat{M}_j$'s, where

$$\hat{M}_j \equiv \hat{\varrho}_t B_j(\mathfrak{F}_{t-1}), \quad j = 1, \dots, J.$$

Let $\hat{S}$ denote the $T \times k$ dimensional score matrix (resulting from a ML estimation) and $\hat{G} \equiv \hat{M}|\hat{S}$ as the stacked $T \times (J + k)$ dimensional matrix of moments and scores, then the CM-test statistic is given by

$$CM \equiv \iota' \hat{G}[\hat{G}'\hat{G}]^{-1}\hat{G}'\iota \overset{a}{\sim} \chi^2_J, \tag{23}$$

with ι a T dimensional vector consisting of ones. In the above formulae, the dependence on the $\sqrt{T}$-consistently estimated parameter vector $\hat{\theta}_0$ is suppressed for notational clarity.

Choose the vector of Borel-functions $B(\mathfrak{F}_{t-1}) \equiv (B_j(\mathfrak{F}_{t-1}))_{j=1,\ldots,J}$ to be

$$B(\mathfrak{F}_{t-1}) = (Y_{t-1}, Y_{t-1}^2, Y_{t-1}^3, Y_{t-2}, Y_{t-2}^2, Y_{t-2}^3, \\ \varepsilon_{t-1}^Y, (\varepsilon_{t-1}^Y)^2, (\varepsilon_{t-1}^Y)^3, \varepsilon_{t-2}^Y, (\varepsilon_{t-2}^Y)^2, (\varepsilon_{t-2}^Y)^3)'. \quad (24)$$

Integrated Conditional-Moment-Test

Bierens (1990) and Jong (1996)

Approximation of the set $\{B(\mathfrak{F}_{t-1})\}$, via a set of exponential functions given by $\{\exp(s \cdot \mathfrak{F}_{t-1})|s \in \mathbb{R}\}$ (same notational convention as above)

The test statistic is designed with s being integrated out. The null hypothesis in the ICM-test has to be changed from (22) to

$$\mathrm{E}[\varrho_t|\mathfrak{F}_{t-1}] = 0 \quad \mathbf{a.s.}, \tag{25}$$

since we allow for exceptions on Lebesgue zero sets.

Density-Forecast

Diebold, Gunther Tay (1998) and Bauwen, Giot, Grammig & Veredas (2000)

Let $\{f^{Y_t}(y_t|\mathfrak{F}_{t-1})\}_{t=1,\dots,T}$ denote the sequence of one step ahead density forecasts, which is the conditional density of Y_t given the past information $\mathfrak{F}_{t-1}$ in an ACD model and $\{p^{Y_t}(y_t|\mathfrak{F}_{t-1})\}_{t=1,\dots,T}$ the sequence of true densities assumed to generate the data generating process. What one actually wants to test, is the equality

$$\{f^{Y_t}(y_t|\mathfrak{F}_{t-1})\}_{t=1,\dots,T} = \{p^{Y_t}(y_t|\mathfrak{F}_{t-1})\}_{t=1,\dots,T}.$$

As $p^{Y_t}(y_t|\mathfrak{F}_{t-1})$ cannot be observed, we use the probability integral transformation of Rosenblatt (1952) , which states that

$$z_t = \int_{-\infty}^{y_t} f^{Y_t}(\nu)d\nu \qquad t = 1, \dots, T \tag{26}$$

is an i.i.d. uniform sequence, under the null hypothesis that $\{f^{Y_t}(y_t|\mathfrak{F}_{t-1})\}_{t=1,\dots,T} = \{p^{Y_t}(y_t|\mathfrak{F}_{t-1})\}_{t=1,\dots,T}$. This hypothesis can be tested for example with the χ^2-independence test against the uniform distribution and/or with e.g. the Ljung-Box- and BDS-test against the i.i.d.ness.

D-Test

Ait-Sahalia (1996) and Fernandes & Grammig (2000)

The idea behind the D-test is to design a test statistic, that measures the proximity between the parametric and the nonparametric estimates of the density functions of the residuals. The null hypothesis is, that there exists a parameter vector such that the parametric density chosen in the ML estimation coincides with the true density of the error terms, i.e.

$$\exists \theta_0 \in \Theta : \quad f_1^{Y_t}(\cdot, \theta_0) = f_1^{Y_t}(\cdot),$$

where Θ denotes the parameter space and $f_1^{Y_t}(\cdot, \theta_0)$ the parametric density of the error terms and $f_1^{Y_t}(\cdot)$ the true density of the error terms. Fernandes and Grammig (2000) propose the following distance that measures the proximity

$$\Psi_{f_1^{Y_t}} = \int_0^\infty \mathbf{1}_{\{\nu \in \mathcal{S}\}} \left(f_1^{Y_t}(\nu, \theta_0) - f_1^{Y_t}(\nu) \right)^2 f_1^{Y_t}(\nu) d\nu,$$

which has the following sample counterpart for the estimated residuals

$$\Psi_{\hat{f}_1^{Y_t}} = \sum_{t=1}^{T} \mathbf{1}_{\{\hat{\varepsilon}_t^Y \in \mathcal{S}\}} \left(f_1^{Y_t}(\hat{\varepsilon}_t^Y, \hat{\theta}) - \widehat{f}_1^{Y_t}(\hat{\varepsilon}_t^Y) \right)^2 ,$$

where $\hat{\theta}$ and $\widehat{f_1^{Y_t}}(\cdot)$ are consistent estimates of θ_0 and $f_1^{Y_t}(\cdot)$ and $\mathcal{S}$ is a compact set. The greater $\Psi_{\hat{f}_1^{Y_t}}$ the less likely is the null hypothesis. Then, under the null and some regularity conditions, the test statistic

$$\hat{\tau}_T^D \equiv \frac{T h_T^{0.5} \Psi_{\hat{f}_1^{Y_t}} - h_T^{-0.5} \hat{\delta}_D}{\hat{\sigma}_D} \xrightarrow{d} N(0,1),$$

where h_T denotes the bandwidth in the nonparametric density estimator. $\hat{\delta}_D$ and $\hat{\sigma}_D$ are consistent estimates of $\delta_D = e_K \, \mathrm{E}(\mathbf{1}_{\{\nu \in \mathcal{S}\}} f_1^{Y_t}(\nu))$ and $\sigma_D^2 = v_K \, \mathrm{E}(\mathbf{1}_{\{\nu \in \mathcal{S}\}} (f_1^{Y_t}(\nu))^3)$ with $e_K = \int_u K^2(u) du$, $v_K = \int_v \left(\int_u K(u) K(u+v) du \right)^2 dv$ and $K(\cdot)$ the kernel function.

Readings:

- Bauwens, L. & N. Hautsch (2009): *Modelling Financial High Frequency Data Using Point Processes*, in: Handbook of Financial Time Series, T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch (eds.), Springer, 953-979.
- Engle, R. F. & J. R. Russell (1998): *Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data*, Econometrica, 66, 1127-1162.
- Engle (2000): *The Econometrics of Ultra-High-Frequency Data*, Econometrica, 68, 1, 1-22.
- Hautsch (2004): *Modelling Irregularly Spaced Financial Data – Theory and Practice of Dynamic Duration Models*, Lecture Notes in Economics and Mathematical Systems, Vol. 539, Springer-Verlag, Berlin.
- Hawkes, A. G. (1971): *Spectra of some self-exciting and mutually exciting point processes*, Biometrika, 58, 83-90.
- Jasiak, J. (1999): Persistence in Intertrade Durations, Working Paper, York University.
- Karr, A. F. (1991): Point Processes and their Statistical Inference, Dekker, New York.
- Russell, J. R. (1999): *Econometric Modelling of Multivariate Irregularly-Spaced High Frequency Data*, Working Paper, University of Chicago.