Modeling bank default intensity in the USA using autoregressive duration models

Siakoulis, Vasilios

21 May 2015

Noname manuscript No.
(will be inserted by the editor)

Modeling bank default intensity in the USA using autoregressive duration models.

Siakoulis G. Vasileios

21/5/2015

Abstract This paper employs a duration based approach in order to model the inter-arrival times of bank failures in the US banking system for the period 1934 - 2014. Conditional duration models that allow duration between bank failures to depend linearly or nonlinearly on its past history are estimated and evaluated. We find evidence of strong persistence along with non-monotonic hazard rates which imply a financial contagion pattern according to which, a high frequency of bank failures generates turbulence which shortly after leads to additional fails, whereas prolonged periods without abnormal events signify the absence of contagious dependence which increases the relative periods between bank failure appearance. In addition, we find that mean duration levels of tranquility spells or equivalently the bank fail events intensity is subject to long run shifts. Further, we obtain statistical significant results when we allow duration to depend linearly on past information variables that capture systemic bank crisis factors along with stock and bond market effects.

Keywords Autoregressive Conditional Duration · Bank Failures · Financial Contagion · Structural breaks
JEL-Classification · C22 · C41 · G01 · G12 · G14

1 Introduction

A bank failure is an event which disturbs the economic environment of a society to a great extend. The depositors are starting to concern themselves about the safety of their savings, the economic policy of a government is set under serious doubt by the public, business lines are cut off, the business activity faces a slowdown and generally the economic environment is destabilized. Throughout economic history, economic crises which were accompanied from a number of bank failures are recorded from the 17th century and the Tulip crisis till the present times.

Atticabank SA (Siakoulis.Vasilios@atticabank.gr). I would also like to acknowledge the contribution of Dr. Venetis Ioannis, professor of econometrics in University of Patras, concerning the understanding and implementation of the methodologies employed in the current work.

The purpose of this study is to propose a new time series approach for the prediction of bank probability of default, which takes into account the existence of contagious dependence, and through that to explore the reasons of a bank failure and examine whether the probability of bank failure may be predicted or not. We adopt the frequently used definition of contagion in the literature as a largely unpredictable, higher correlation after a major distress event (see Pesaran and Pick, 2004). In the domain of bank failures contagion means that there are fluctuations in the rate of fails that cannot be explained by common factors. Focardi and Fabozzi (2005) expressed contagious defaults behaving in this way as point processes whose intensity is conditional upon previous intervals between defaults through a self-exciting ACD point process. In this framework, a high frequency of bank failures generates turbulence in the banking sector which shortly after leads to additional fails, whereas prolonged periods without abnormal events signify the absence of contagious dependence which increases the relative periods between bank failure appearance.

As such, we construct a duration series that measures the inter-arrival times of bank failures in the USA. Strictly speaking a bank is considered to fail when the regulatory authorities decide that it must stop its operation due to its inability to respond in the demand of deposits from its depositors. Therefore the failure of a bank is a legal and not an economic event since a bank fails when supervisors say and not when it has presented severe financial problems. This policy may change the timing of bank failures relatively to economic events. If supervisors close banks only after they had a chance to work out their problems then there will be a lag between economic crises and banking crises (Nuxoll, 2003). In order to resolve this issue we also classify as failed the banks which are found in need of government assistance.

The generated distress events are irregularly spaced in time and the inter-arrival times of bank failures exhibit clustering behavior and autoregressive dynamics similar to that of GARCH models (Engle, 1982 and Bollerslev, 1986). Since we are dealing with a duration series, we employ the Autoregressive Conditional Duration (ACD) model of Engle and Russel (1998) to model the inter-arrival times of bank failure events. To the best of our knowledge, such an approach has not been attempted in the past in the domain of contagious bank failures while it is one of few applications of autoregressive conditional duration models outside the high frequency transactions data literature (e.g. Christoffersen and Pelletier (2004), Fischer and Zurlinden (2004), Hamidieh et al. (2013)).

Also this approach could add to the toolbox relative to corporate probability of default prediction in line with the propositions of Focardi and Fabozzi (2005). Giesecke et al. (2010) using an extensive data set spanning the period from 1866 to 2008 showed that corporate bond market has repeatedly suffered clustered default events. In our case by employing an autoregressive conditional duration framework combined with a set of explanatory variables we depart from the standard doubly stochastic model of default, under which cross-firm default correlation that is associated with observable factors determining conditional default probabilities accounts for all the degree of time clustering of defaults. This doubly-stochastic assumption is overly restrictive to the extent that default of one firm could have an important direct influence on the default of another firm (Duffie et al. (2007)). Motivated by the results of Das et al. (2007) who find that a standard doubly stochastic model of firm default intensities fails to fully capture the clustering of defaults of U.S. industrial firms between 1979 and 2004, Giesecke et al. (2008) propose a default timing model in which the default intensity evolves through time according to a self-exciting model. According to them two important effects, frailty and contagion generate self-exciting phenomena, so that a firm default tends to increase the likelihood of other firms to fail also. In the first case a default reveals information about unobservable (frailty) factors and therefore has an influence on the conditional default rates of any other firms that depend on the same factors, whereas contagion is based on the propagation of financial distress through contractual linkages among firms. Duffie et al (2009) find strong evidence for the presence of unobservable risk factors and they estimate that they have a large impact on fitted conditional mean default rates, above and beyond those predicted by observable default covariates.

Our contribution lies in modeling the correlation structure of inter-arrival times of bank faiures. Within self-contained autoregressive models, we find significant evidence of clustering behavior and duration persistence. The duration process exhibits self-exciting behavior, i.e., demonstrates a high degree of persistence, even when conditioned on variables related to various risk factors. The degree of persistence does not diminishes even when low frequency variation in the conditional duration mean is encompassed in the form of mean level shifts. Bank crisis effects along with influences from bond and stock markets are present, affecting significantly the probability of bank failure occurence.

The rest of the paper is structured as follows. In section 2, we provide an overview of the alternative ACD specifications used in the paper. Section 3 contains a summary of the methodologies employed for evaluating density forecasts. Section 4 contains the data description and the empirical application. In section 5, we introduce additional explicative variables in the conditional mean duration specification. Section 6 tests for the presence of unconditional duration mean level shifts and investigates their potential effect on estimated persistence. Finally, section 7 offers some concluding remarks.

2 Autoregressive Conditional Duration model (ACD)

Let $\{t_0, t_1, ...\}$ be a stochastic process of strictly increasing arrival times with $0 = t_0 < t_1 < ... < t_{N(t)} < ... < t_{N(T)} = n$ associated with a counting function $N(t)$, the number of events that have occurred up to time $t \in [0, T]$. A fundamental application of the ACD model by Engle and Russel (1998) is its ability to parametrically measure and forecast the intensity of arrivals or duration between two events that occur at times t_{i-1} and t_i denoted by $x_i = t_i - t_{i-1}$, $i = 1, ..., n$.

The ACD model assumes that

$$x_i = E(x_i|\Omega_{i-1}; \theta_x)\, \varepsilon_i$$

where Ω_{i-1} denotes the information set available at time t_{i-1} and includes at least the observed values $\{x_1, x_2, ..., x_{n-1}\}$ and ε_i is non-negative i.i.d with density $f(\varepsilon; \theta_\varepsilon)$. Setting $\psi_i = E(x_i|\Omega_{i-1}; \theta_x)$ the model implies that $E(\varepsilon_i) = 1$ and all the temporal dependence of the duration process x_i is captured by the conditional expected duration.

The benchmark ACD model proposed by Engle and Russel (1998) is built on two parameterizations. The first assumes a GARCH-like structure for the conditional expected duration

$$\psi_i = \omega + \alpha_1 x_{i-1} + \beta_1 \psi_{i-1}$$

while the second attaches the exponential density function with mean 1 to ε_i.

The multiplicative error structure of the model implies that the log-likelihood function is given by

$$\ln L(\theta) = \sum_{i=1}^{n} l_t(\theta) = \sum_{i=1}^{n} \left[\ln f\left(\frac{x_i}{\psi_i}; \theta_\varepsilon\right) - \ln(\psi_i) \right] \tag{1}$$

where $\theta = \left(\theta_x', \theta_\varepsilon'\right)'$ is the $k \times 1$ vector of all the unknown parameters and in the case of benchmark ACD model $\theta_x = (\omega, \alpha_1, \beta_1)'$ while θ_ε is empty.

The ACD model, similarly to the GARCH models for conditional heteroskedasticity, is very general and allows a variety of duration conditional mean specifications and distributions for the multiplicative innovation ε_i. Hence, the benchmark model, where ε_i is assumed non-negative i.i.d with exponential density $f_E(\varepsilon_i; \theta_\varepsilon) = e^{-\varepsilon_i}$, has been extended in the literature in two aspects.

First, the conditional mean specification is parametrically enriched to account for empirically valid asymmetries to past information. We employ the augmented ACD (AACD) model of Fernandes and Grammig (2006) that nests most of the parametric specifications proposed in the literature and is given by

$$\psi_i^\lambda = \omega + \alpha_1 \psi_{i-1}^\lambda \left[|\varepsilon_{i-1} - b| + c(\varepsilon_{i-1} - b) \right]^v + \beta_1 \psi_{i-1}^\lambda \tag{2}$$

The power transformation of ψ_i along with the extended parametric introduction $\left[|\varepsilon_{i-1} - b| + c(\varepsilon_{i-1} - b) \right]^v$ of past standardized durations in the conditional mean, allows a wide variety of shapes for the shock impact function[1]. The shift parameter $b > 0$ allows for a kinked impact curve and sign asymmetries with positive (negative) shock impact when $\varepsilon_{i-1} > (<) b$. The rotation parameter c allows for clockwise (counter-clockwise) $c > (<) 0$ rotation of $|\varepsilon_{i-1} - b|$ implying less (more) sensitivity to shocks that are $\varepsilon_{i-1} > b$. Parameters λ and v play a similar role, introducing concavity or convexity in the response implying decreasing or increasing conditional duration sensitivity with respect to shock magnitudes.

Second, more flexible error distributions are adopted that generalize the benchmark exponential distribution for random variables with positive support. Grammig and Maurer (2000) show that quasi maximum likelihood estimation of θ based on (1) can be biased in finite samples that can be as large as 15,000 observations. A direct consequence of biased estimates is the reduced predictive performance of the ACD models. Moreover, the exponential distribution implies a flat hazard rate function that measures the transition rate from a duration spell of a certain length given that the spell lasted until that moment, assigning equal probability to long and short durations. Such an assumption can be empirically invalid.

Engle and Russel (1998) extend the flat hazard rate case to monotonic increasing or decreasing hazard rate functions by assuming that ε_i are i.i.d following the

[1] The curve that traces the impact of a shock ε_{i-1} on the conditional mean duration ψ_i for a given value of ψ_{i-1} and the remaining parameters.

Weibull distribution with parameter γ, thus $\theta_\varepsilon = \gamma$. If $\gamma = 1$, then the Weibull ACD model reduces to the benchmark exponential ACD model. The Weibull density function is given by

$$f_W(\varepsilon_i;\theta_\varepsilon) = \gamma\left(\Gamma\left(1+\frac{1}{\gamma}\right)\right)^{\gamma}\varepsilon_i^{\gamma-1}\exp\left\{-\left(\Gamma\left(1+\frac{1}{\gamma}\right)\varepsilon_i\right)^{\gamma}\right\}$$

with hazard rate function[2] admitting the form

$$h_W(\varepsilon_i;\theta_\varepsilon) = \gamma\varepsilon_i^{\gamma-1}$$

If $\gamma > 1$ ($\gamma < 1$) the hazard rate function is increasing (decreasing), that is long durations will be more (less) likely.

Still, monotonicity can be restrictive in empirical applications. Grammig and Maurer (2000) proposed the use of the Burr distribution with density function

$$f_B(\varepsilon_i;\theta_\varepsilon) = \frac{\kappa\mu^{\kappa}\varepsilon_i^{\kappa-1}}{\left(1+\gamma\mu^{\kappa}\varepsilon_i^{\kappa}\right)^{1+\frac{1}{\gamma}}}$$

where $\theta_\varepsilon = (\kappa,\gamma)$, $\kappa > \gamma > 0$ with $\kappa > m\gamma$ ensuring that the m^{th} moment exists while

$$\mu = \frac{\Gamma\left(1+\frac{1}{\kappa}\right)\Gamma\left(\frac{1}{\gamma}-\frac{1}{\kappa}\right)}{\gamma^{1+\frac{1}{\kappa}}\Gamma\left(1+\frac{1}{\gamma}\right)}$$

The cumulative distribution function is given by

$$F_B(\varepsilon_i;\theta_\varepsilon) = 1-\left(1+\gamma\mu^{\kappa}\varepsilon_i^{\kappa}\right)^{-\frac{1}{\gamma}}$$

and the hazard function by

$$h_B(\varepsilon_i;\theta_\varepsilon) = \frac{\kappa\mu^{\kappa}\varepsilon_i^{\kappa-1}}{1+\gamma\mu^{\kappa}\varepsilon_i^{\kappa}}$$

The Burr distribution[3] nests both the Weibull distribution as $\gamma \to 0$, and the exponential distribution as $\gamma \to 0, \kappa \to 1$ as well as the log-logistic distribution when $\gamma \to 1$. The Burr distribution allows for general non-monotonic hazard functions under suitable parameter restrictions, for example empirically valid U-shape or inverted U-shape hazard functions where short and long durations, in terms of large deviations of x_i from its conditional mean, are progressively less likely to occur.

Finally, Allen et al. (2008), Allen et al. (2009) and Sun et al. (2008) propose the use of the log-normal distribution with density function

$$f_{LN}(\varepsilon_i;\theta_\varepsilon) = \frac{1}{\sigma\sqrt{2\pi}\varepsilon_i}\exp\left\{-\frac{\left(\ln\varepsilon_i+\frac{\sigma^2}{2}\right)^2}{2\sigma^2}\right\} \tag{3}$$

[2] The hazard rate function of the multiplicative error term is named baseline hazard.

[3] Lunde (1999) and other authors consider the generalized gamma distribution which is also a two shape parameter distribution than can deliver rich non-monotonic hazard function shapes. The generalized gamma distribution also contains the exponential and Weibull distributions as special cases.

that corresponds to the density of the exponent of a normally distributed random variable with variance σ^2 and mean $-\frac{\sigma^2}{2}$. The value of the mean is chosen in order to guarantee that the log-normal terms ε_i have expected value one. The hazard function is given by

$$h_{LN}(\varepsilon_i; \theta_\varepsilon) = \frac{f_{LN}(\varepsilon_i; \theta_\varepsilon)}{1 - \Phi\left(\frac{\ln \varepsilon_i + \frac{\sigma^2}{2}}{\sigma}\right)}$$

where Φ is the cumulative density function of a standard normal random variable. The log-normal density (3) also produces various hazard function shapes, however, it has only one free parameter compared to the two free parameters of the Burr distribution (and other similar distributions), hence, it might well be the preferred choice for empirical applications.

The estimation of all ACD models is performed by maximum likelihood using the BFGS algorithm and suitable initial conditions. Asymptotic standard errors are based on the outer-product-of-the-gradient (OPG) estimator of the information matrix since the augmented model nonlinearities render Hessian-based (and sandwich form) estimates difficult to compute due to numerical problems.

3 Evaluating the ACD model

Model evaluation is based on conditional mean specification and on the distribution assumption regarding standardized durations. To compare for in-sample overparameterization we compute the AIC and BIC criteria (BIC penalizes more heavily parameter addition). In addition, Ljung-Box statistics are calculated for the standardized $\hat{\varepsilon}_i = \frac{x_i}{\psi_i}$ and squared standardized $\hat{\varepsilon}_i^2$ durations as preliminary descriptives on remaining linear autocorrelation. Detailed model specification will be based on a general to specific modeling approach given the parametric structure of the models considered and the discussion in Fernandes and Grammig (2006, p.16) that if one fails to start from a sufficiently general specification may be directed to misleading outcomes since there are various parameter combinations that can work interchangeably.

Another way to test both the conditional mean and distributional assumption of the models is to evaluate the density forecasts of alternative models both in-sample and out-of sample. For this purpose, we use the method developed by Diebold et al. (1998) to test the forecasting performance of nested and non-nested general dynamic models. Bauwens et al. (2004) used extensively the method to compare different ACD specifications. The method consists in computing the sequence of empirical probability integral transforms (PIT), $z_i = F\left(\hat{\varepsilon}_i; \hat{\theta}_\varepsilon\right)$, $i = 1, ..., n$, that is, the cumulative probability of the observed standardized durations $\hat{\varepsilon}_i$ under the assumed forecast distribution. If the underlying model specification is accurate, then z_i will be i.i.d uniformly distributed $U(0, 1)$ and will have no-autocorrelation left neither in level nor when raised to integer powers. In order to assess how close the distribution of z_i is to a uniform distribution, Diebold et al. (1998) propose the use of intuitive graphical methods and also a simple inspection of the correlograms for the series $(z_i - \bar{z})^p$, $p = 1, ..., 4$ to examine for autocorrelation and potentially more sophisticated nonlinear forms of dependence. Hence, we visually compare the estimated density of z_i to that of a $U(0, 1)$ using a histogram plot and approximate bin height confidence intervals under the null hypothesis of i.i.d. $U(0,1)$. In addition, we visually inspect quantile-quantile (QQ) plots where the closer the plot is to the 45 degrees line the closer the distribution of z_i is to the uniform distribution. All correlograms are examined on the basis of the usual Bartlett confidence intervals.

The Diebold et al. (1998) graphical inspection methods are informative with respect to the correct conditional calibration of density forecasts. However, it is difficult to test for uniformity in small samples. We complement the model evaluation procedures by using the Berkowitz (2001) likelihood ratio (LR) tests that are based on the transformed variable $u_i = \Phi^{-1}(z_i)$ where Φ^{-1} is the inverse of the standard normal distribution. Under the null hypothesis of accurate model specification, the transformed variables u_i are distributed as i.i.d $N(0,1)$ which allows maximum likelihood estimation and facilitates the construction of LR tests with good finite sample properties. In particular, an unrestricted $AR(1)$ specification for u_i is estimated and a group of likelihood ratio statistics are calculated for the null hypothesis, $H_0^{(1)}: \mu = 0$, $H_0^{(2)}: \sigma^2 = 1$, $H_0^{(3)}: \rho = 0$, $H_0^{(4)}: \mu = 0$, $\sigma^2 = 1$, $H_0^{(5)}: \mu = 0, \sigma^2 = 1, \rho = 0$. These test statistics are termed $LR^{(1)}$, $LR^{(2)}$, $LR^{(3)}$, $LR^{(4)}$, $LR^{(5)}$ respectively and are distributed as χ^2 random variables with degrees of freedom equal to the number of restrictions. Given that we test for mean, variance and autocorrelation deviations from the null, the LR test has power against general alternatives that accommodate misspecification in the conditional mean, variance and dynamics of the model in use.

4 Empirical Application

Our dataset comprises of the dates in which US banks failed, or were found in need of government assistance, provided from the Federal Deposit Insurance Coorporation (FDIC)[4], spanning the period from 2/4/1934 to 25/2/2014. Next, we construct a binary distress event variable that takes the value of 1 at the days where at least one US bank failed or found in need of assistance as shown in Fig. 1 top panel. Those days denote the $t_i, i = 1, ..., n$ arrival times sequence. Finally, we construct the dependent duration variable $x_i = t_i - t_{i-1}$ as the number of days between two distress events. The duration variable is shown in Fig. 1 bottom panel, with $n = 1466$ available sample observations.

Fig. 1 here

Note that this duration variable does not take zero values since when bank failures occur for two or more consecutive days in a row, the next period is considered to start from the last crisis day.

The autoregressive form of ACD-type models allow us to capture duration clustering, that is, long (short) durations followed by long (short) durations between failure events. A long duration, namely the time elapsed between two bank fails, marks tranquil periods whereas a sequence of short durations marks periods of financial turbulence where banks tend to collapse one after the other either due to economic factors or contagious banks crises.

[4] https://www.fdic.gov/bank/individual/failed/banklist.html

Flexible distribution forms for the standardized durations ε_i allow various hazard function shapes. In our application, the hazard rate $h(\varepsilon;\theta_\varepsilon)$, where ε is excess duration (observed over conditional mean duration), measures the instantaneous rate of change from a tranquil state to a state where a bank failure occurs given that no bank failures have occurred to date. Alternatively, it measures the "probability"[5] of bank default. For example the value $h(1;\theta_\varepsilon)$ would measure the probability of bank default given that no bank defaults have appeared for a period equal to the conditional mean of the series, whereas $h(2;\theta_\varepsilon)$ would measure the probability of exit from a longer tranquility period of two times the conditional mean given that no bank defaults have appeared during that period.

In Table 1, we report the coefficient estimates[6], together with the corresponding p-values of the ACD models described in Section 2 and in-sample model fit measures described in section 3.

table 1 here

In terms of the log-likelihood value and the AIC and BIC criteria, the augmented specification (2) outperforms marginaly the benchmark ACD model. In turn, the flexible distribution approach that uses the Burr or lognormal distributions outperforms the exponential and Weibull models.

The Ljung-Box statistics on the standardized durations and the squared standardized durations do not show any signs of remaining autocorrelation. Also the Berkowitz (2001) LR tests confirm the superiority of the Burr and lognormal distributions.

Fig. 2 contains the z_i histograms with 20 bins for all estimated models. There is discernible difference in the histogram shape for models under alternative error distributions. Deviations from uniformity are evident for the exponential and Weibull-ACD that find difficult to account for the durations at the lower bound of the distribution. It seems that they systematically produce biased in-sample fits for very small durations. The z-histograms show clearly the superiority of the Burr model. The AACD Burr specification produces empirical integral transforms that match the implied theoretical density very well and tends to give accurate in-sample fits over the whole range of observed values although a spike is visible for small durations (third bin). Also, the AACD Burr histogram is visibly smoother than its ACD counterpart. Given that the exponential and Weibull distribution models will not further considered as the data do not support either flat or monotonic increasing/decreasing hazard rate functions.

Fig. 2 here

In the ACD case the sum of the coefficients $a_1 + \beta_1$, that expresses the autoregressive part in the ARMA representation of the duration series, is greater than 0.95 in the majority of models indicating a very high degree of persistence. In the AACD case, the autoregressive parameter β_1 is smaller than 1, a feature

[5] Notice that the hazard rate is not necessarily bounded from above by one.

[6] All empirical work, including the figures, was conducted in R 2.14.0. Maximum likelihood routines were based on package *maxLik* while structural break tests were based on package *strucchange*. We have also employed packages *sandwich* and *portres*.

that guarantees stationarity of the process at least in the Box-Cox version[7] of the models that fits the data best, as it will become apparent.

The Burr parameter estimates $\hat{k}$ and $\hat{\gamma}$ are quite robust regardless of the conditional mean specification. The ratio $\hat{k}/\hat{\gamma} = 4.126$ with (delta approximation) standard error 0.5296 so that the upper bound of the approximate 95% confidence interval for k/γ is 5.16.

The coefficient estimates in table 1 further reveal the necessity of the double Box–Cox transformation $\lambda \neq v$ in the AACD models. There is strong evidence supporting convergence of λ to zero (i.e. support for the log-transformation), whereas v is bounded away from 0 in both the Burr and lognormal specifications. The fact that the estimated a_1 in the AACD specifications is statistically insignificant is explained by the close to zero values of $\hat{\lambda}$.

In light of $\lambda \to 0$ and $v = 0.623$ it is not surprising that the shift and rotation parameter estimates b, c are insignificant. Fernandes and Grammig (2006) show that letting λ free to vary and accounting for asymmetric effects seem to operate as substitute to the introduction of asymmetric responses in specifications with fixed λ. Thus, the preferred in-sample model reduces to the Boc-Cox ACD specification with the Burr density being the most successful distribution candidate.

The maximum likelihood estimates of the preferred model are shown below

$$\ln \psi_i = \underset{(0.052)}{-0.170} + \underset{(0.054)}{0.204}\,\varepsilon_{i-1}^{\overset{0.623}{(0.137)}} + \underset{(0.004)}{0.992}\ln \psi_{i-1}\ ,\ \hat{\kappa} = \underset{(0.060)}{1.563}\ ,\ \hat{\gamma} = \underset{(0.063)}{0.378} \tag{4}$$

$$\ln \hat{L} = -4681.42\ ,\ AIC = 6.395\ ,\ BIC = 6.416\ ,\ Q(4) = 0.6140$$

Under the logarithmic transformation II and when $v = 1$ [8], although durations have a linear effect on the log conditional mean, the news impact curve is not linear but convex, so that the sensitivity of ψ_i to shocks in ε_{i-1} is lower if ε_{i-1} is small and higher if ε_{i-1} is large. But in our case where $v = 0.623$. the adjustment process of the conditional mean to recent durations follows an almost linear pattern as it can be seen from a plot of the empirical shock impact curve

$$SIC_i = \exp\left\{-0.170 + 0.204\varepsilon^{0.623} + 0.992\ln\left(\bar{\hat{\psi}}\right)\right\}$$

for shock sizes $\varepsilon = \{0.05, 0.1, ..., 14\}$ which is given in the top panel of Fig. 3.

fig. 3 here

A plot of the parametric estimate for the baseline[9] hazard function is given in Fig. 3 bottom panel. It shows a non-monotonic inverted and asymetric U shape indicating the distress event arrival asymmetries, hence, the distress event arrival intensity is low for very short and very long durations. The rate at which very low ($\varepsilon << 1$) and very long ($\varepsilon >> 1$), relative to the conditional mean, tranquility

[7] See Fernades and Gramming (2006) for detailed analysis. The Box-Cox version of AACD model (2) is obtained when $\lambda \to 0$. and b = c = 0.

[8] This model is referred to Fernandez and Grammig (2006) as Logarithmic ACD Type II

[9] Often, the conditional hazard function

$$h\left(x_i|\Omega_{i-1}\right) = h\left(\varepsilon_i\right)\frac{1}{\psi_i}$$

is cited. The results where not quantitatively or qualitatively different.

spells will be completed, given that they last until that moment, is small which is intuitively consistent with duration clustering. The hazard function achieves a maximum at around $\varepsilon = 1.31$. Given that the average sample conditional mean duration is approximately 17 days, the maximum hazard is attained for 22.3 days long tranquility spells.

Finally, for purposes of comparison, Fig. 4 contains the histogram for the z_i sequence, QQ-plots and correlograms for the z_i and u_i sequences implied by the preferred model.

fig. 4 here

The third bin spike persists despite the parsimonious and better fit of the model implying that small (but not very small) durations are over-represented. The correlogram for both the PIT and Berkowitz transformations does not reveal remaining linear dependence. However, the correlograms for the squared processes, not reported for brevity, present some spikes over the confidence intervals suggesting some heteroskedasticity or nonlinear features that remain hidden.

We complete the univariate time series specification of the duration series by executing a forecasting experiment. Bauwens and Hautsch, (2009, p.964) propose out-of-sample evaluations of alternative ACD specifications, in order to avoid the problem of potential over-fitting. Dufour and Engle (2000) find that ACD models perform better than autoregressions in durations or log durations, however, the forecasts are generally not very good although they offer improvements in out-of-sample results. In addition, they find that the distribution choice does not significantly affect the short-run conditional mean forecasts but it has measurable effects on density forecasts. Similar conclusions were reached by Bauwens et al. (2004) where simple model approaches were found to perform as well as more complex model specifications.

Our out-of-sample experiment is based on truncating the sample endpoint at $n_1 < n$, estimating sequentially the conditional duration models using subsamples of size $n_1, n_1+1, ..., n-1$ and forecasting one step ahead. The results, for $n_1 = 1200$ are summarized in table 2.

table 2 here

We adopt a number of forecasting measures as in Dufour and Engle (2000). Table 2 reveals that the gains between alternative models are not impressive when forecasting out-of-sampling the conditional mean, as with previously mentioned studies.

5 Enhancing the past information set

The preceding time series analysis is self-contained. One may wish to include additional predictive variables to enhance the forecasting power of the model or to test theory related hypotheses by including predetermined variables related to bond yield determination. See, Engle and Russell (1998), Bauwens and Giot (2000, 2003) and Zhang et al. (2001), among others, for the inclusion of predetermined variables in market microstructure related ACD models.

Hautsch (2012) presents two types of additive introduction of a $g \times 1$ vector of explanatory variables $z_i = (z_{1,i}, ..., z_{g,i})'$ in an ACD model. These are, the dynamic type

$$\ln \psi_i = \omega + \alpha_1 \varepsilon^v_{i-1} + \beta_1 \ln \psi_{i-1} + \lambda' z_{i-1} \quad (5)$$

and the static type

$$\begin{aligned} \left(\ln \psi_i - \lambda' z_{i-1}\right) &= \omega + \alpha_1 \varepsilon^v_{i-1} + \beta_1 \left(\ln \psi_{i-1} - \lambda' z_{i-2}\right) \Rightarrow \\ \ln \psi_i &= \omega + \alpha_1 \varepsilon^v_{i-1} + \beta_1 \ln \psi_{i-1} + \lambda' z_{i-1} + \lambda^{*\prime} z_{i-2} \end{aligned} \quad (6)$$

written here in terms of the best univariate model specification chosen in the previous section. In (5), the effects of z_{i-1} carry over to ψ_i according to an infinite lag structure and $|\beta_1| < 1$ guarantees the exponential decay of the effects on distant future durations. In (6), the $1 \times g$ parameter vector $\lambda^{*\prime}$ satisfies the restrictions $\lambda^{*\prime} = \beta_1 \lambda'$. The effects are static with z_{i-1} affecting only ψ_i through λ and this type of model is more suitable whenever the regressors are connected to certain time periods, for example dummy variables that would capture one-off history events. Further explanatory variables can also be introduced multiplicatively as a scaling function, however such an assumption is used predominately to "deseasonalize" in-day trade durations and is not applicable in our model. In addition, the multiplicative introduction of the regressors would imply that the conditional mean parameters are varying with z_i, an assumption not readily justified.

Concerning the selection of explanatory variables we receive guidance from the studies of Allen and Gale (1998) which pointed the effects of financial contagion in the banking sector along with Mishkin and White (2002) which noted the resilience of banking sector with financial markets. In the last case stock market effects affect the health of the banking system due to adverse selection problems in credit markets, moral hazard issues and depositors panic. Also turbulent periods in debt markets may also provoke phenomena of bank panic as in the case of sub-prime crisis where banks and depositors couldn't determine the amount of exposure to MBS (Mortgage Bank Securities) of their counterparties leading to lack of confidence and drying-off liquidity in the interbank markets (Demyanyk and Hasan, 2010). In this framework we examine the hypotheses that bank crisis effects along with influences from stock and bond markets affect significantly the probability of bank failure occurence.

As such, we consider the inclusion of number of failed bank, in the day preceding each tranquil period, as factor indicative of bank crisis pressure and financial contagion effects. More precisely we distinguish the case where the distress event occured due to one bank failure only (BANKNUM = 1) and the case where the event was caused by two or more banks collapsing in the same day (BANKNUM = 2) indicating a systemic shock effect [10].

In order to capture stock market conditions we construct a binary variable (FCRISE) taking the value of one in the months where a fall beyond the 5 per cent quantile of the distribution occured for the monthly returns of the S & P index [11].

[10] The data were found from the Federal Deposit Insurance Coorporation (FDIC) https://www.fdic.gov/bank/individual/failed/banklist.html

[11] The S & P index data span the period from Jan. 1871 till Apr. 2014 and were taken from the dataset used in Robert J. Shiller's book, Irrational Exuberance [Princeton University Press 2000, Broadway Books 2001, 2nd ed., 2005] are available for download available at http://www.econ.yale.edu/ shiller/data.htm

Relatively to the bond market distress event indicators we use the difference between the Moody's seasoned corporate BAA bond yield and the Moody's seasoned corporate AAA bond yield (CSPREAD). Monthly data for corporate AAA and BAA bond yields were taken from the FRED database of the Federal Reserve Bank of Saint Louis [12]. As recorded in Rigobon (2002) evidence of money market turbulence is a flight to quality effect where the spreads over high credit quality bonds on liquid exposures (Aaa bonds) narrow while at the same time spreads on Baa bonds substantially widen.

Estimates of model (5) with the predetermined variables are reported in Table 3.

table 3 here

A systemic shock in the banking sector as expressed through the simultaneous collapse of more that one commercial banks decreases the expected inter-arrival time and increases the probability of a subsiquent bank collapse. This is evident from the from the statistically significant (p-value equals 0.0023) and negative coefficient estimate of BANKNUM. In line with theory shocks in the bond markets (CSPREAD) have negative effects reducing the expected inter-arrival time till the next bank failure which is interpreted as an increase in crisis probability in banking sector (hazard rate). However, financial shocks in the stock markets are not significant (p-value equals 0.2329) even though, in line with theory, the coefficient FCRISE appears negative.

The log likelihood value of (5) shows an improvement over model (4) while the histogram for the z_i sequence, QQ-plots and correlograms for the z_i and u_i sequences implied by model (5) mark a marginal improvement.

6 Structural breaks in mean duration levels

A remarkably consistent feature of all estimated models is the high degree of persistence as captured by the estimated autoregressive coefficient $\hat{\beta}_1$ or by the sum $\hat{\alpha}_1 + \hat{\beta}_1$. Even when explanatory variables are introduced, the sum remains above 0.9. In the conditional heteroskedasticity literature, large part of the observed persistence in high frequency volatility estimates is attributed to structural breaks, see for example Perron and Qu 2010. Zhang et al. (2001) question how plausible the stationarity assumption is with transactions duration data and they find strong evidence of parameter breaks closely aligned with real economic events.

Given the abovementioned, we proceed to test for abrupt shifts (breaks) in the mean duration level but we do not allow for changing dynamics through time periods. To estimate the number of breaks and the location of the break points, we implement an information criteria approach using the BIC criterion that requires searching for the global minimum of the residual sum of squares employing the efficient search algorithm of Bai and Perron (2003). In order to guard against under-estimation of the number of breaks when non-monotonic mean shift reversions are present, we use a sequential search method coupled with repartition as in Bai (1997). The search for breaks is applied on the log duration $\ln(x_i)$ series. The

[12] http://research.stlouisfed.org/fred2/

log transformation produces series "closer" to Gaussianity and stabilizes the variance of the series by reducing the noise levels around potential mean level shifts, hence, facilitates better statistical results from the breaks search procedure.

We are able to identify three breaks in the mean level of (log) duration at sample points $\{\hat{\imath}_1, \hat{\imath}_2, \hat{\imath}_3,\} = \{246, 465, 1192\}$ that correspond to the following dates 18/9/1940, 14/3/1981 and 2/5/1993. We then split the sample period into 4 subsample regimes and estimate the BC-ACD model

$$\ln \psi_i = \omega + \alpha_1 \varepsilon_{i-1} + \beta_1 \ln \psi_{i-1}^{v} + \lambda' z_{i-1} + \sum_{j=1}^{3} \delta_j D_j \tag{7}$$

with dummy variables constructed as $D_j = 1\left\{i > \hat{\imath}_j\right\}$, $j = 1, ..., 3$. Parameters $\delta_1, ..., \delta_3$ show the marginal change in mean duration levels between subsequent regimes. Maximum likelihood estimates of the model with the abovementioned explanatory covariates and the mean shift dummies are reported in Table 4 while Fig. 5 shows the observed duration series along with the estimated conditional duration from model (7).

table 4 here

A number of remarks can be made,

Remark 1 all dummy variable parameter estimates show high statistical significance and point to large differences across segments, in particular, a relative increase of 55.5% after 18/9/1940 (since $\frac{\hat{\delta}_1}{\hat{\omega}} = 0.555$), a decrease of -67.9% after 14/3/1981 and an increase of 45.6% in the last sub-sample after 2/5/1993.

Remark 2 even though structural brakes are introduced in the specification the persistence remains as measured in the Box-Cox case by the conditional autoregressive coefficient β_1.

Remark 3 contrary to the model 5, with the introduction of mean shift dummies the Ljung-Box statistics at 4 lags for levels and squares of the residuals do not reject the null hypothesis. From this perspective only when we include mean shifts there is no remaining unexplained structure in the residuals. However few spikes remain in a few lags when we examine the histogram of the empirical probability integral transforms z_i as well as correlograms for the series $(z_i - \bar{z})^p$, $(u_i - \bar{u})^p$, $p = 1, ..., 4$. In any case all other specification diagnostics have accomplished well.

Remark 4 when the mean shift dummies are included in the model, the stock market crisis variable FCRISE is significant with a p-value around 0.10. Thus the information carried by the variable is significant when long run shifts are introduced in the inter-arrival times of bank fail event durations.

fig. 5 here

As a robustness test to alternative sample periods and to assess the impact of the global financial crisis, we restrict the sample to the pre-crisis period ending in 2006 and we re-estimate the models of tables 1, 4 and 5 for the subsample period April 1934 to December 2006. All results are available from the author. The available duration observations were reduced to 1277. The autoregressive parameters estimates $\hat{\omega}$, $\hat{\beta}_1$ and distribution parameter estimates $\hat{k}$ and $\hat{\gamma}$ in table 5 remain virtually unchanged. The parameter λ increases from 0.75 to 1.44, altering the shape of the empirical shock impact curve from linear to exponential. The structural break findings are also the same with minor discrepancies relatively to break date locations. Finally, the impact of the explanatory variables remains qualitatively unchanged albeit the exclusion of the relatively small inter-arrival durations for bank failures events in the financial crisis episode has a decreasing effect on the parameter estimates BANKNUM and CSPREAD whereas the FCRISE effect becomes statisticaly insignificant. To sum up, the robustness check reveals that the main results are unaffected by the exclusion of the financial crisis (and post-financial crisis) period.

7 Concluding remarks

We have applied autoregressive conditional duration models to a duration series that measures the inter-arrival times of bank failures in the USA. With respect to the econometric challenge of modelling the conditional duration series, our results confirm the findings of Fernandes and Grammig (2006), Bauwens et al. (2004) and Allen et al. (2009), with the Burr and log-normal distributions offering superior performance both in-sample and out-of-sample in comparison to the exponential and Weibull distributions. The Burr distribution is particularly well suited for the series at hand given its parametric simplicity and its ability to model small durations. The produced bank fail arrival intensity is low for very short and very long durations. In addition, evidence is consistent with the self-exciting hypothesis where the no-failure event duration demonstrates a high degree of persistence irrespective of the conditioning on predetermined variables or structural breaks that might affect the time duration of distress events. A systemic bank crisis effect, an effect from financial markets extreme negative movements and an explicative variable linked to turbulence in bond markets constitute significant factors for the duration process.

8 Appendix

Fig. 1 Top panel: binary distress event variable, bottom panel: duration variable $x_i = t_i - t_{i-1}$ measures the number of days between two distress events

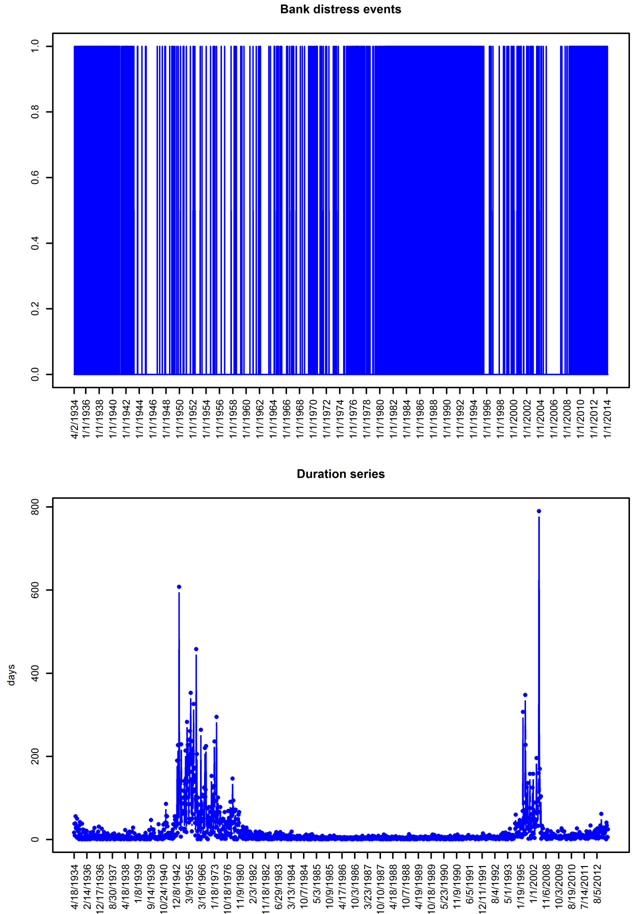

Fig. 2 The flat lines superimposed on all histograms are approximate 95% confidence intervals for the individual bin heights under the null hypothesis that the empirical integral transforms z_i of the residuals by each estimated model are distributed as $i.i.d.U(0,1)$.

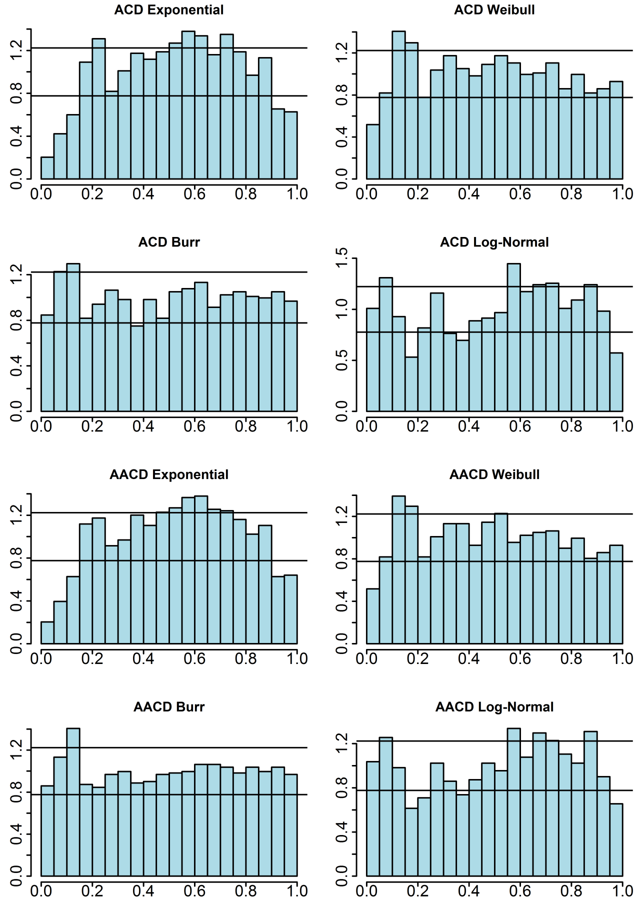

Fig. 3 Empirical shock impact curve (top panel) and baseline hazard function (bottom panel) derived from model (4) parametric estimates.

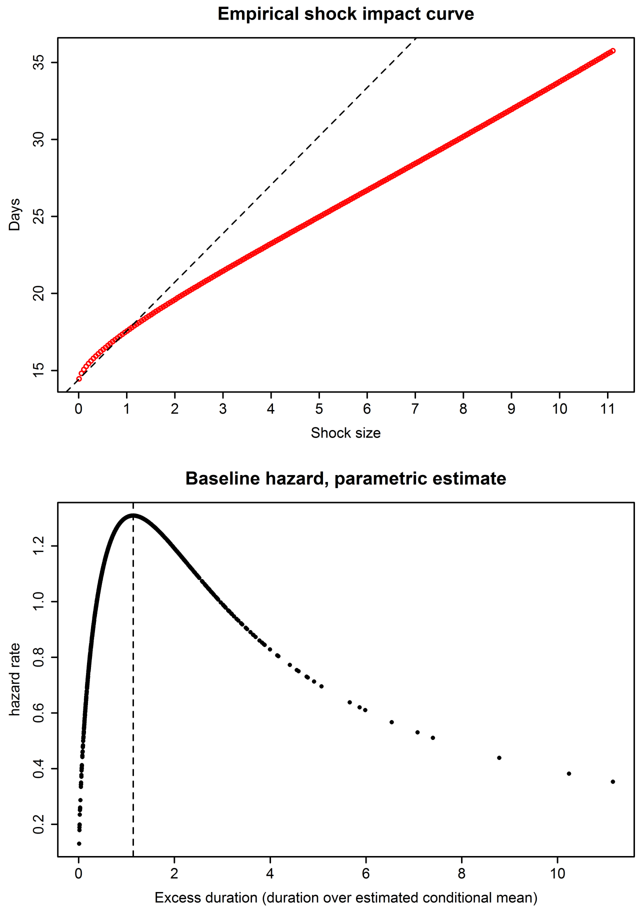

Fig. 4 Graphical methods results for model (4). The flat lines superimposed on all correlograms are Bartlett's approximate 95% confidence intervals.

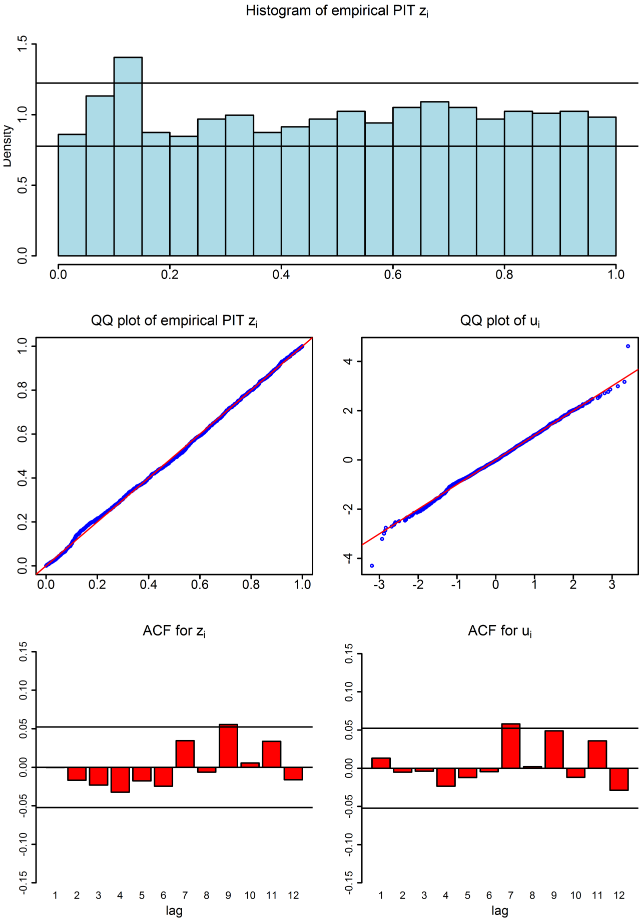

Fig. 5 Sub-sample regimes with distinct mean duration level.

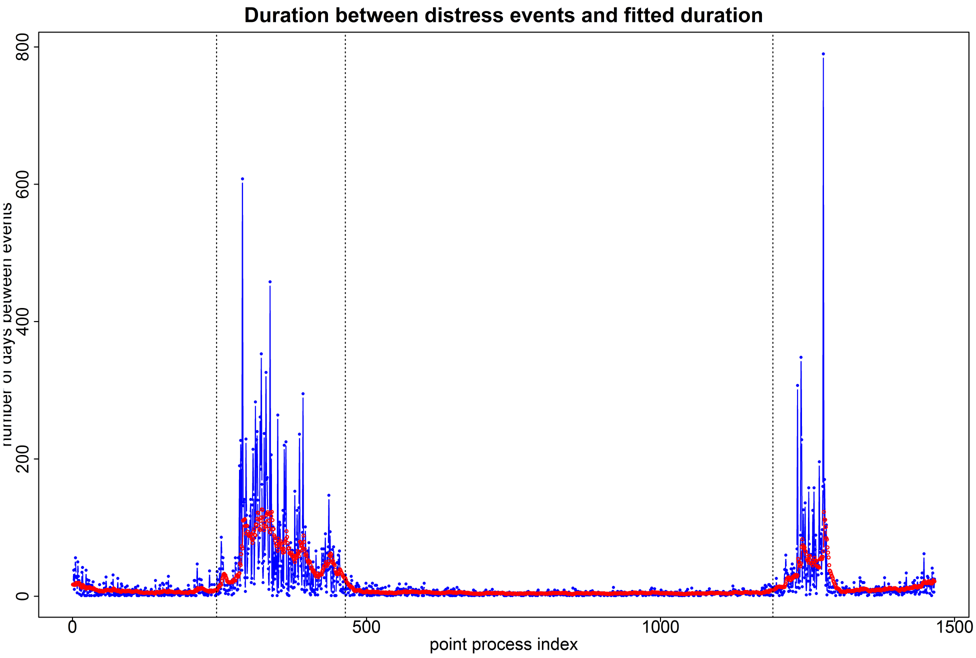

Table 1 Estimation results for the ACD and augmented ACD (AACD) models under alternative error distributions -E: exponential, -W: Weibull, -B: Burr, -L: Lognormal

	ACD-E	ACD-W	ACD-B	ACD-L	AACD-E	AACD-W	AACD-B	AACD-L	BCACD-B
omega	0.0727	0.0364	0.1491	0.3762	0.0051	0.0020	0.0088	0.0294	-0.1702
p-val.	0.186	0.337	0.000	0.000	0.867	0.874	0.706	0.433	0.001
a	0.1559	0.1742	0.1351	0.1434	0.0463	0.0573	0.0179	0.0412	0.2045
p-val.	0.000	0.000	0.000	0.000	0.965	1.000	1.000	0.828	0.000
b	0.8416	0.8415	0.8498	0.8210	0.9757	0.9771	0.9870	0.9607	0.9916
p-val.	0.000	0.000	0.000	0.000	0.008	0.009	0.000	0.000	0.000
v					0.7089	0.7438	0.6247	0.6372	0.6230
p-val.					0.106	0.252	0.028	0.000	0.000
lamda					0.1055	0.1102	0.0255	0.0817	
p-val.					0.957	0.957	0.977	0.852	
c					-0.6500	-0.6862	-0.8617	-0.7521	
p-val.					0.805	0.999	0.999	0.125	
beta					0.0153	0.0067	0.0085	0.0205	
p-val.					0.812	0.992	0.979	0.569	
k			1.5591				1.5633		1.5629
p-val.			0.000				0.000		0.000
gamma		1.2284	0.3740			1.2291	0.3789		0.3780
p-val.		0.000	0.000			0.000	0.000		0.000
sigma				0.9140				0.9127	
p-val.				0.000				0.000	
lnL	-4767.77	-4717.72	-4682.57	-4715.53	-4767.45	-4717.51	-4681.69	-4713.46	-4681.42
AIC	6.509	6.442	6.395	6.439	6.514	6.447	6.399	6.441	6.395
BIC	6.519	6.456	6.413	6.453	6.539	6.476	6.432	6.470	6.416
Q(4)	0.7841	0.6689	0.8118	0.2224	0.8561	0.7200	0.6188	0.0631	0.6140
Q(4)-sq	0.8957	0.9517	0.5284	0.3177	0.8287	0.9474	0.2193	0.1204	0.2242
LR1	0.0000	0.5065	0.9323	0.9323	0.0000	0.5336	0.9526	0.9746	0.9290
LR2	0.5268	0.7986	0.9833	0.9582	0.5248	0.7967	0.9740	0.9595	0.9807
LR3	0.6913	0.6916	0.5111	0.6375	0.7007	0.7425	0.4859	0.6563	0.4729
LR4	0.0000	0.7765	0.9962	0.9978	0.0000	0.7972	0.9977	0.9992	0.9957
LR5	0.0000	0.8804	0.9320	0.9732	0.0000	0.9042	0.9210	0.9782	0.9957

Notes Figures in parentheses correspond to standard errors based on the outer product gradient (OPG) estimator of the information matrix. lnL reports the value of the log-likelihood function, whereas AIC and BIC denotes the Akaike and Schwarz information criteria. $Q(4)$ and $Q_2(4)$ correspond to the p-values of Ljung–Box statistic for up to 4th order serial correlation in the standardized model residuals and squared standardized model residuals respectively. LR1 to LR5 display the Berkowitz (2001) likelihood ratio tests applied to standardized durations.

Table 2 Out-of-sample 1-step-ahead forecasting results.

	MSE	RMSE	MAE	MAPE	EXP	CORR
ACD-E	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ACD-W	1.0235	1.0117	1.0618	1.1203	0.9951	1.0000
ACD-B	0.9910	0.9955	0.9702	0.9425	1.0157	0.9965
ACD-L	0.9849	0.9924	0.9393	0.8709	1.0062	1.0066
AACD-E	0.9893	0.9946	0.9933	0.9956	0.9945	1.0200
AACD-W	1.0183	1.0091	1.0581	1.1120	0.9935	1.0081
AACD-B	0.9766	0.9882	0.9563	0.9324	1.0073	1.0259
AACD-L	0.9625	0.9811	0.9166	0.8740	0.9931	1.0704
BCACD-B	0.9716	0.9857	0.9453	0.9335	1.0012	1.0358

Notes All measures are divided by the measure of ACD-E. MSE: mean square error, RMSE: root MSE, MAE: mean absolute error, MAPE: mean absolute percentage error, EXP: exponential loss function measure, CORR: measures correlation between actual and forecasted values. See Dufour and Engle (2000) for details on all measures.

Table 3 Estimates of the final BCACD specification with Burr density and predetermined variables.

	omega	**a**	**b**	**v**	**k**	**gamma**	**BANKNUM**	**CSPREAD**	**FCRISE**
Param	-0.0277	0.1783	0.9827	0.6078	1.5602	0.3575	-0.0508	-0.0195	-0.0251
(se)	0.0710	0.0707	0.0051	0.2164	0.0562	0.0628	0.0167	0.0067	0.0211
P-val	0.6962	0.0117	0.0000	0.0050	0.0000	0.0000	0.0023	0.0038	0.2329
Model statistics									
lnL	**AIC**	**BIC**	**Q(4)**	**$Q\hat{2}(4)$**					
-4671.67	6.3856	6.4181	0.5841	0.0951					
LR1	**LR2**	**LR3**	**LR4**	**LR5**					
0.9190	0.9831	0.2326	0.9951	0.6973					

Table 4 Estimates of the final BCACD specification with Burr density, predetermined variables and structural breaks.

	omega	**a**	**b**	**v**	**k**	**gamma**	**BANKNUM**	**CSPREAD**	**FCRISE**
Param	0.1164	0.1171	0.9511	0.7548	1.5726	0.3578	-0.0670	-0.0234	-0.0373
(se)	0.0636	0.0490	0.0109	0.2466	0.0571	0.0639	0.0195	0.0078	0.0229
P-val	0.0673	0.0169	0.0000	0.0022	0.0000	0.0000	0.0006	0.0028	0.1041

	d1	**d2**	**d3**
Param	0.0645	-0.0790	0.0530
(se)	0.0213	0.0226	0.0135
P-val	0.0025	0.0005	0.0001

Model statistics

lnL	**AIC**	**BIC**	**Q(4)**	**Q$\hat{2}$(4)**
-4659.11	6.3726	6.4159	0.7561	0.7257

LR1	**LR2**	**LR3**	**LR4**	**LR5**
0.8949	0.9812	0.3178	0.9910	0.7972

References

Allen, F., and Gale, D., 2000. Financial Contagion . Journal of political economy, 108, 1

Allen, D., Chan, F., McAleer, M., Peiris, S., 2008. Finite sample properties of the QMLE for the Log-ACD model Application to Australian Stocks. Journal of Econometrics 147, 163-185

Allen, D., Lazarov, Z., McAleer, M., Peiris, S., 2009. Comparison of Alternative ACD Models via Density and Interval Forecasts: Evidence from the Australian Stock Market. Mathematics and Computers in Simulation 79, 2535-2555

Bai, J., 1997. Estimating multiple breaks one at a time. Econometric Theory 13, 315-352

Bai, J., Perron, P., 2003. Computation and Analysis of Multiple Structural Change Models. Journal of Applied Econometrics 18, 1-22

Bauwens, L., Giot, P., 2000. The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks, Annales d'Economie et de Statistique 60, 117-149

Bauwens, L., Giot, P., Grammig, J., Veredas, D., 2004. A comparison of financial duration models via density forecasts. International Journal of Forecasting 20, 589-609

Bauwens, L., Hautsch, N., 2009. Modelling Financial High Frequency Data Using Point Processes, in: Handbook of Financial Time Series, Edited by: Andersen TG, Davis RA, Kreiss JP, Mikosch T, Springer-Verlag Berlin Heidelberg, pages 953-979

Berkowitz, J., 2001, Testing density forecasts, with applications to risk management. Journal of Business & Economic Statistics 19, 4, 465-474

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307-327

Christoffersen, P., and Pelletier, D., 2004. Backtesting value-at-risk: A duration-based approach. Journal of Financial Econometrics 2, 1, 84-108

Das, S., Duffie, D., Kapadia, N., Saita, L., 2007. Common failings: how corporate defaults are correlated. The Journal of Finance 62, 93-117
Demyanyk, Y., Hasan, I., 2010. Financial crises and bank failures: A review of prediction methods. Omega, 315-324.
Diebold, F., Gunther A., Tay, S., 1998. Evaluating density forecasts with applications to financial risk management. International Economic Review 39, 863-883
Duffie, D., Saita, L., Wang, K., 2007. Multi-period corporate default prediction with stochastic covariates. Journal of Financial Economics 83, 3, 635-665
Duffie, D., Eckner, A., Horel, G., Saita, L., 2009. Frailty Correlated Default. The Journal of Finance 64, 5, 2089-2123
Duffur, A., and Engle, R., 2000. The ACD model: Predictability of time between consecutive trades. ICMA Centre Discussion Papers in Finance 2000-05
Engle, RF., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 4, 987-1006
Engle, RF., Russell, JR., 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. Econometrica 66, 1127-1162
Fernandes, M., Grammig, J., 2006. A family of autoregressive conditional duration models. Journal of Econometrics 130, 1-23
Fischer, A., and Zurlinden, M., 2004. Are interventions self-exciting? Open Economies Review 15, 223-237
Giesecke, K., Azizpour, S., Seminar For Discussions, Ding, X., Kim, B., Mudchanatongsuk, S., 2008. Self-Exciting Corporate Defaults: Contagion vs. Frailty (2008). Working Paper, Stanford University
Giesecke, K., Longstaff, A., Schaefer, S., Strebulaev, I., 2010. Corporate bond default risk: A 150 year perspective. NBER Working Paper 15848.
Grammig, J., Maurer, K., 2000. Non-monotonic hazard functions and the Autoregressive Conditional Duration model. Econometrics Journal 3, 16-38
Hamidieh, K., Stoev, S., Michailidis, G., 2013. Intensity-based estimation of extreme loss event probability and value at risk. Applied Stochastic Models in Business and Industry 29, 3, 171-186
Hautsch, N., 2012. Econometrics of Financial High-Frequency Data. Springer-Verlag Berlin Heidelberg
Lunde, A., 1999. A generalized Gamma autoregressive conditional duration model. Working Paper, Department of Economics, Politics and Public Administration, Aalborg University, Denmark.
Mishkin, F., and White, E., 2002. U.S. Stock Market Crashes and Their Aftermath: Implications for Monetary Policy. NBER Working Paper, 8992
Pacurar, M., 2008. Autoregressive conditional duration models in finance: a survey of the theoretical and empirical literature. Journal of Economic Surveys 22, 711-751
Perron, P., Qu, Z., 2010. Long-Memory and Level Shifts in the Volatility of Stock Market Return Indices. Journal of Business & Economic Statistics 28, 2, 275-290
Rigobon, R. 2002. International Financial Contagion: Theory and Evidence in Evolution. The Research Foundation of AIMR Working Paper (No. 08/2002).
Sun, W., Rachev, S., Fabozzi, F., Kalev, P., 2008. Fractals in trade duration: capturing long-range dependence and heavy tailedness in modeling trade duration, Annals of Finance 4, 217-241

Zhang, M., Russell, J., Tsay, R., 2001. A Nonlinear Conditional Autoregressive Duration Model with Applications to Financial Transactions Data. Journal of Econometrics 104, 179-207