sciendo

Scenario based merger & acquisition forecasting

Kainat KHOWAJA
Humboldt-Universität zu Berlin, Berlin, Germany

Danial SAEF
Humboldt-Universität zu Berlin, Berlin, Germany
danial.saef@hu-berlin.de

Sergej SIZOV
Thieme Group, Germany

Wolfgang Karl HÄRDLE
Humboldt-Universität zu Berlin, BRC Blockchain Research Center, Berlin; Sim Kee Boon Institute, Singapore Management University, Singapore; NUS, Center of Competitiveness, Singapore; National Chiao Tung University, Prague, Czech Republic; the European Cooperation in Science & Technology COST Action [CA19130]; Grant CAS: XDA 23020303 and DFG IRTG 1792 gratefully acknowledged

Abstract. *While there is no doubt that M&A activity in the corporate sector follows wave-like patterns, there is no uniquely accepted definition of such a "merger wave" in a time series context. Count-data time series models are often employed to measure M&A activity and merger waves are then defined as clusters of periods with an unusually high number of M&A deals retrospectively. However, the distribution of deals is usually not normal (Gaussian). More recently, different approaches that take into account the time-varying nature of M&A activity have been proposed, but still require the a-priori selection of parameters. We propose adapting the combination of the Local Parametric Approach and Multiplier Bootstrap to a count data setup in order to identify locally homogeneous intervals in the time series of M&A activity. This eliminates the need for manual parameter selection and allows for the generation of accurate forecasts without any manual input.*

Keywords: mergers and acquisitions, forecasting, non stationary time series, Poisson modelling, change point detection.

Please cite the article as follows: Khowaja, K., Saef, D., Sizov, S. and Härdle, W. K. (2024), "Scenario based merger & acquisition forecasting", *Management & Marketing*. Vol. 19, No. 4, pp. 579-600, DOI: 10.2478/mmcks-2024-0026.

Introduction

Predicting M&As remains challenging due to the fluctuations caused by their wave like market behavior (Harford 2005). The resulting non-stationarity makes traditional time series models inapplicable. We propose a method to find the stationary span within the

recent past that can be safely used for making decisions about the near future and show how it can be used on the example of M&A activity in several German industries.

The method in this paper essentially proposes a new way of thinking about the occurrence of M&A events. It differs from previous approaches as it (I) only uses the assumption that there exist stationary intervals within the time series, (II) works even with very small sample sizes, (III) does not require a specification of a sample horizon, or the number and nature of different states, (IV) does not require manual inputs or data screening as it is fully automated, and thus does not introduce a bias or mis-specification through human interaction, (V) the proposed approach is not only a stand-alone method, but also serves as an extension to models like ARIMA: we can identify intervals within the time series where the model assumptions are met. Finally, we apply the procedure to the often overlooked non US markets.

Past research has often focused on M&As US markets exclusively. For example, Very et al. (2012) comprehensively discusses applications of the most common approaches ARIMA and Markov Switching Models on M&As in US Markets. Unfortunately, these approaches can not easily be generalized to non US markets. Consider Figure 1, which depicts the number of mergers and acquisitions in the German energy market. It shows the wave like behavior of M&As and the presence of non-stationarity in such time series. To overcome these problems, consider other business problems that are commonly described by count data models. Typical examples are call center arrivals or supply chains. Instead of using a time series approach, they often simulate from the event distribution that is assumed to be Poisson. Unlike these examples, the wave like behavior in M&A time series however distorts the distribution. Figure 2 illustrates the total distribution, with a yellow line to describe the 95th percentile of the distribution. While the majority of observations are between 0 and 7, and could follow a Poisson distribution, merger waves generate some heavy outliers that follow a different law.

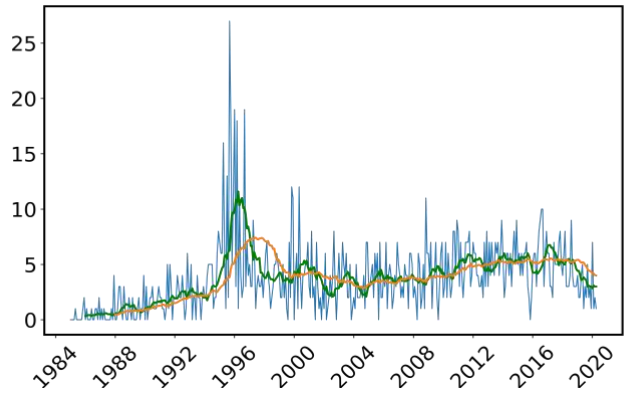


Figure 1. Time series of count of mergers and acquisitions per month/energy sector (GER), with moving average curve of 1 year and 3 years

Source: Authors' own research.

Bianchi and Chiarella (2018) underlines the importance of accounting for the timevarying dynamics of M&As, but their method requires careful specification of the underlying parameters, which usually change over time. This problem can be avoided by using a subset with stable windows to separate different periods of activity. We take

advantage of the fact that within a time series there are always intervals where the parameter doesn't change. Hence, we can detect locally homogeneous time windows within the time series where a stable distribution can be assumed. However, such selection is often done by hand, as it is challenging to describe the different states using statistical methods. Such approaches are neither scalable, nor adaptive.

Look again at Figure 1. It indicates that selecting a stable window by hand is not always optimal. Taking small intervals like one year for parameterizing the time series shows high variance (instability), whereas taking longer intervals like 3 years can introduce a high bias due to the time varying nature of the parameter. While traditional time series analysis methods like Bayesian Changepoint Detection (BCP) are commonly employed for identifying shifts in data patterns and selecting stable windows, they often necessitate manual parameterization, introducing potential biases in the process.

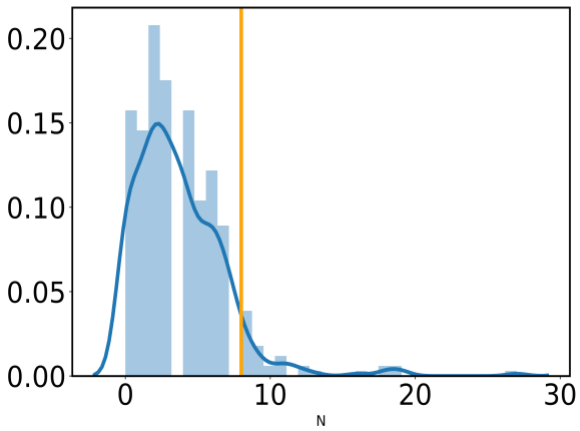


Figure 2. Distribution of count of mergers and acquisitions per month for the energy sector (GER), and tail of distribution (95th percentile) indicated by yellow line

Source: Authors' own research.

Furthermore, these methods may not always strike an optimal balance between variance and bias. This is because traditional change point detection algorithms are designed with the goal of identifying all change points with maximum probability. However, when it comes to window selection, the detection of the last change point in the time series is crucial. Therefore, it becomes imperative to revise the algorithm to optimize the probability of accurately identifying the last change point, which is pivotal for effective window selection.

Our method allows to balance the trade-off between variance and bias by providing a practical technique that chooses the appropriate interval within the time series. It thus allows us to forecast the evolution of mergers & acquisitions accurately in a fully automated fashion. Our results show that the procedure generates reliable forecasts in the studied markets (telecommunication, energy, financials), and reveal that the dynamics of these industries have distinct characteristics that need careful consideration during model selection.

We employ an adaptive estimation method called Local Parametric Approach (LPA) following Spokoiny (2009) to account for the heterogeneous nature of the different industries, and to deem manual parameterization unnecessary. It helps us find locally homogeneous time intervals with stable parameters and guarantees a trade-off between parameter variance and modeling bias. The technique is based on a series of likelihood ratio

tests to determine assumed but unknown change points in the underlying series. As a result one finds local intervals of homogeneity and efficient estimates at each point in time.

Since we are often dealing with small sample sizes, and the test statistic used in our method has an unknown distribution, we need a method to approximate this distribution. Recent advances in bootstrapping methodology allow us to generate confidence sets and critical values that non-asymptotically approximate the true distribution. Here, we couple LPA with multiplier bootstrap (MBS) (Spokoiny and Zhilova 2015) for approximating a critical value for the testing procedure.

Comparing the forecast accuracy to commonly used methods shows that our model can generate accurate forecasts without the need for parameterization, or any manual input. Additionally, we highlight that even though ARIMA models have high accuracy, the best fit ARIMA model frequently changes within different various intervals of the time series. In terms of forecast accuracy measures, our approach is more accurate than using simpler metrics such as a 12 or 36 month moving averages or even benchmark models like BCP in the financial and telecommunication industries. Only the energy was better approximated using moving averages. However, even in such cases our approach could be used as a guideline for selecting an appropriate interval making forecasts with ARIMA models. This is more accurate than using moving averages while maintaining a balance in the bias-variance trade-off. Re-thinking the way that M&A events occur helps us in making better predictions by using novel statistical methods and thus helping managers and controllers to make better decisions using data analytics driven insights.

The remainder of this paper is structured as follows: Section discusses related work. Section describes the algorithm that is based on a combination of LPA, MBS and put into action in an iterative procedure. Section 1 contains an experimental study. We describe the evaluation method, verify the robustness of the presented algorithm in simulation scenarios, apply it empirically on a dataset of mergers & acquisitions and show how it can be used to generate forecasts. Section 2 presents key results, such as robustness to diverse data inputs and adaptability to other applications. Section 3 discusses limitations, such as in the evaluation approach or computational aspects and suggests next steps like density forecasting, introducing a judgemental component and extending the test-statistic.

Literature review

Our approach extends previous literature as it serves as a generic toolbox that could easily be adapted to other applications and can give accurate forecasts that can, but do not have to be, adjusted by incorporating knowledge from industry experts and is adaptable to arbitrary frequencies. Although we show how to forecast the occurrence of M&A events, our methodology could also extend research in other areas that typically use Poisson processes such as demand forecasting for supply chain planning (Yelland, Kim, and Stratulate 2010), call center arrival times (Taylor 2007, 2011) and Oreshkin, Réegnard, and L'Ecuyer (2016), sales forecasting (Kolsarici and Vakratsas 2015), or mergers & acquisitions forecasting (Very et al. 2012; Akkus, Cookson, and Hortaçsu 2015; Bianchi and Chiarella 2018). However, the available data sets are usually subject to non-stationarity and structural breaks, and they usually make manual efforts to fit a meaningful model necessary.

Empirical evidence strongly suggests that mergers often have "wave" like clusters in time, see Martynova and Renneboog (2005), Harford (2005) and Maksimovic, Phillips, and Yang (2013). Ahern and Harford (2014) find that such an activity is subject to network effects

and that these "waves" largely occur within industries, but can also be transmitted to connected industries. Furthermore, shocks of any kind, even if they lead to merger waves, are difficult to predict. "Wave" patterns seem to be heterogeneous and differ both in time and in industries. Following their argumentation, we conclude that data on M&A should be evaluated per industry and geographic location due to differences in regulation, innovation power, technology, and stock markets.

Predictive models for M&A intensity could be specified through aggregating acquired knowledge and tailored to specific industries and markets. Alternatively, time-series models, e.g. ARMA can be used. While a co-variate based model provides explainability to the user, it requires manual efforts to gather data and incorporate expert knowledge to define relevant variables and calibrate their impact on a predictive model. Data gathering can be time-consuming and expensive, and knowledge about modeling heterogeneous industries in a specific application requires domain knowledge, which is often scarce and narrowed to said application. Time series models are an alternative, as they can be adapted to any other data set with comparable structure. However, such models need to be robust to non-stationarity, structural breaks and wave patterns that limit their predictive power.

There is no doubt that only time varying time series models approximate the dynamics of the underlying series better than any homogeneous parameter approach, e.g. a fixed ARMA(p,q) model. Therefore, we employ an adaptive estimation method called Local Parametric Approach (LPA). The quantitative implementation was first proposed in Spokoiny (1998), advances are made in Mercurio and Spokoiny (2004) and Spokoiny (2009). It helps us to find locally homogeneous time intervals with stable parameters and guarantees a trade-off between parameter variance and modeling bias. The technique is based on a series of likelihood ratio tests to determine some distributional change in the underlying series (which in our case is the maximum likelihood estimator of the Poisson process). As a result one finds local intervals of homogeneity and efficient estimates at each point in time.

Since we are often dealing with small sample sizes, and the test statistic distribution is unknown, we need a method to approximate this distribution. Recent advances in bootstrapping methodology allow us to generate empirically approximate the distribution and hence the critical values that non-asymptotically approximate the true distribution. Here, we couple LPA with multiplier bootstrap (MBS) (Spokoiny and Zhilova 2015) for approximating a critical value for the testing procedure. MBS builds up on wild bootstrap, that originates from Wu (1986) and Beran (1986). Klochkov, Härdle, and Xu (2019) also presents theoretical justifications for why MBS is appropriate to be coupled with LPA, even though for a different setting. An important application is reported in Härdle and Mammen (1993), and Mammen (1993). Further advances are made in Chatterjee and Bose (2005) and Arlot, Blanchard, and Roquain (2010). Notable publications that precede Spokoiny and Zhilova (2015) are Bücher and Dette (2013) and Chernozhukov, Chetverikov, and Kato (2013). Klochkov, Härdle, and Xu (2019) presents an application in the context of a conditional autoregressive Value at Risk model. We generalize this LPA idea to any count data that is locally Poisson distributed, although our simulation study indicates that this assumption could be relaxed to the general membership to any of the exponential families. That allows us to bridge business requirements with the robustness of novel statistical methods through flexible automated estimations. This is valuable as available data for problems in strategic management like forecasting the intensity of M&A events is often

challenging. Assumptions on the underlying distribution, such as stationarity or the absence of structural breaks are usually not fulfilled.

We detect locally homogeneous windows by computing a non-parametric likelihood ratio statistics. This approach can be linked to the branch of change point detection methods. Chen and Gupta (2011), Eckley, Fearnhead, and Killick (2011), and Aminikhanghahi and Cook (2017) summarize and evaluate diverse methods of change point detection. Notable approaches are Hinkley and Hinkley (1970), (Hsu 1979), (Haccou, Meelis, and van de Geer 1987), and (Chen and Gupta 1999) that propose change point detection methods for gamma, exponentially, and normally distributed data respectively. Kutoyants and Spokoiny (1999) propose an adaptive procedure as well as theoretical properties for Poisson distributed data. Chen and Gupta (2011) provides the null distribution of a likelihood ratio test for Poisson distributed random variables, but they do not evaluate the efficiency of the procedure on real data.

Both change point detection and homogeneous window approaches can serve as determinator for an optimal forecasting window. Recent approaches for optimal window selection are Giraitis, Kapetanios, and Price (2013), Pesaran, Pick, and Pranovich (2013), and Inoue, Jin, and Rossi (2017). They address parameter instability and frequent structural breaks and conclude that adaptive window selection should be preferred over choosing fixed window sizes, such as a 1 year or 3 year moving average. Since we aim at developing a generally applicable method that is robust to different data characteristics, we need an adaptive forecasting window. Hence, we pursue a non-parametric approach that is independent of knowledge about or assumptions on the data set, except that the values are generated by a Poisson process with smooth but time varying intensity over some unknown time window.

Note that even though change point detection methods can be modified to deal with the problem of finding homogeneous windows within a time series, LPA is not a change-point detection method, at least not in the conventional manner. The aim of LPA is to look back from a given point in time, and find another time point where the distribution changes. This change can be characterized as the change in mean for some distributions, but not necessarily so. In contrast, traditional change point detection methods usually detect anomalous sequences/states in a time series by looking at the whole time series. Many times, these methods aim to find multiple points within the time series to represent temporary shocks or structural breaks that may or may not have caused a change in distribution. Due to the difference between our approach and change point detection methods, we will not discuss or compare our approach with change point detection methods.

Methodology

Basic idea

Planning processes in corporate environments are based on internal financial data of different kinds. Traditionally, these problems have been solved using diverse time series models. However, it is difficult to use them since real time series often have structural breaks and they do not satisfy stationarity assumptions. Following the same reasoning, fitting a single parametric model to a non-stationary time series is not appropriate. To contribute to solving this problem, we focus on detecting locally homogeneous intervals with stable parameters within a ”stability region” of a time series. In these so called ”intervals of homogeneity”, a local parametric model can be safely used. To be more specific, we focus on

a count data model where the time varying intensity follows a Poisson process. Take again figure 1 as an example. We aim to detect the last time point (possibly around 95-97) that added the non stationary component (a slight uptrend is observable) and cut off the time series before the point in time where the parameter changes. We call these time points break points in the rest of this paper. Automated analyses can be beneficial as they make modelling easier. Our procedure automatically detects the locally homogeneous region and use only this time window for analysis and forecasting.

In this section, we lay down the procedure of finding locally homogeneous windows, and show how it can be used to obtain forecasts.

Stochastics

Let $Y_t \in \mathrm{N}, t = 0,...,T$ be a count data time series such as the count of M&A series, see in Figure 1. Think of $Y_t \sim Poisson(\theta)$, where θ represents the rate or average number of occurrences in a fixed interval. Since we allow for time variation in our model, for any interval $I = [a,b]$ with $a < b$ and $a,b \in \{0,...,T\}$, we write $(Y_t)_{t\in I} \sim Poisson(\theta)$.

The log likelihood function on I is:

$$L_I(\theta) = \sum_{t\in I} \log(\theta^{Y_t} e^{-\theta}/Y_t!) \quad = \log\theta \sum_{t\in I} Y_t - \sum_{t\in I} \theta - \sum_{t\in I} \log(Y_t!), \tag{1}$$

and the MLE $\widetilde{\theta_I}$ based on observations in $i \in I$ is:

$$\tilde{\theta}_I \stackrel{\text{def}}{=} \underset{\theta\in\Theta}{\operatorname{argmax}}\ L_I(\theta), \tag{2}$$

which for a Poisson model is the sample mean.

Local Parametric Approach

LPA, first introduced by Spokoiny (1998) is based on the phenomenon that a series of locally parametric models can describe the features of a time series better than a global parametric model. The basic idea is that given a time series and a model for its dynamics, one finds locally stationary intervals of the time series in an iterative fashion. This is done by finding the set of the most recent observations, such that the model parameters are approximately stable in that interval. This set of time points is called *interval of homogeneity*. Employing the same procedure at each point in time, one locally estimates the parameter (Härdle, Hautsch, and Mihoci 2015). The merit of LPA is that it does not require an explicit expression of the law of the dynamics of the parameter, but only assumes that the parameter is constant on some unknown time interval in the past (Spokoiny 2009).

In order to check the homogeneity of an interval $I = [a,b]$, LPA looks for some break point $\tau \in (a,b)$ such that $A_\tau = [a,\tau)$ has a parameter and $B_\tau = [\tau,b]$ has a parameter as well. Obviously if the parameters are identical, we are looking at a homogeneous interval I. Deviations from the identical parameter situation indicate then non-homogeneity. One then checks this homogeneity for different points τ in the center of I (Klochkov, Härdle, and Xu 2019).

The testing hypotheses are therefore:

$$\begin{aligned} &H_0(I) : (Y_t)_{t\in I} \sim Poisson(\theta^*_I), \theta^*_I \in \Theta, \\ &\qquad \text{vs} \\ &H_1(I) : (Y_t)_{t\in A_\tau} \sim Poisson(\theta^*_{A_\tau}), \theta^*_{A_\tau} \in \Theta, \\ &\qquad (Y_t)_{t\in B_\tau} \sim Poisson(\theta^*_{B_\tau}), \theta^*_{B_\tau} \in \Theta, \\ &\qquad \text{with some } \theta^*_{A_\tau} \neq \theta^*_{B_\tau}, \end{aligned} \tag{3}$$

and the LR test statistic for a break point τ is:

$$\mathcal{T}_{I,\tau} = L_{A_\tau}(\tilde{\theta}_{A_\tau}) + L_{B_\tau}(\tilde{\theta}_{B_\tau}) - L_I(\tilde{\theta}_I), \quad (4)$$

which, since one has many candidates $\tau \in J$, arrives at:

$$\mathcal{T}_I = \max_{\tau \in J} \mathcal{T}_{I,\tau}, \quad (5)$$

this unfortunately has a very intractable distribution, hence the critical values $\mathfrak{z}_I(\alpha)$ needed to test H_0 in equation (3)

$$\mathcal{T}_I \geq \mathfrak{z}_I(\alpha) \quad (6)$$

are hard to calculate. Indeed, the limiting distribution of T_1 is different from general likelihood ratio tests due to the nonlinear max operator in equation (5). Hence, convergence of the generalized LR statistics to a χ^2 distribution according to Wilk's phenomenon can not be put into action. While the asymptotic distribution of the sup-LR test in equation (5) can still be derived (Andrews and Ploberger 1994), a large enough sample size is required for its asymptotic critical values to be applicable. Certainly, that is not the case in most practical situations where only small samples of data are available.

Spokoiny and Zhilova (2015) provide a non-asymptotic result for mis-specified models with small sampling sizes. They propose a simulation based approach called multiplier bootstrap (MBS) which can be used to construct critical values. We extend the MBS technique and discuss it in detail in the next section.

Multiplier bootstrap

Since the distribution for T_1 (5) is not available, we employ bootstrapping. Multiplier bootstrap is a novel and state-of-the-art technique to approximate the unknown distribution. It does not put any assumptions on the data distribution and constructs the critical values completely in a data driven fashion. The departure from distributional assumptions makes this technique unique and easily applicable with LPA to assist in hypothesis testing. Refer to Spokoiny and Zhilova (2015) for the theoretical justification of using the multiplier bootstrap procedure for the construction of likelihood-based confidence sets.

The idea of MBS is to first introduce random weights to the previously defined likelihood function:

$$L_I^\circ(\theta) = \sum_{t \in I} w_t l_t(\theta),$$

where the weights w_t are iid, independent of the sample, with $E(w_t)=1$ and $Var(w_t)=1$. The weights generated under these assumptions are optimal for mimicing the unknown distribution (see assumption 3.1 in Klochkov, Härdle, and Xu (2019))). The bootstrap version of (1) is given by:

$$L_I^\circ(\theta) = \log\ \theta \sum_{t \in I} Y_t w_t - \sum_{t \in I} w_t \theta - \sum_{t \in I} \log(Y_t!) w_t.$$

The bootstrap MLE is then defined as:

$$\tilde{\theta}_I^\circ = \arg\max L_I^\circ(\theta),$$

which is:

$$\tilde{\theta}_I^\circ = \frac{\sum_{t \in I} Y_t w_t}{\sum_{t \in I} w_t}.$$

It follows that the corresponding bootstrap of (8) is:

$$\mathcal{T}_{I,\tau}^\circ = L_{A,\tau}^\circ(\tilde{\theta}_{A,\tau}^\circ) + L_{B,\tau}^\circ(\tilde{\theta}_{B,\tau}^\circ) - \sup_\theta \left\{ L_{A,\tau}^\circ(\theta) + L_{B,\tau}^\circ(\theta + \tilde{\theta}_{B,\tau} - \tilde{\theta}_{A,\tau}) \right\}. \quad (7)$$

Here, the shift term $\widetilde{\theta_{B,\tau}} - \widetilde{\theta_{A,\tau}}$ is devoted to compensate for the biases of the bootstrap estimators $\widetilde{\theta_{A,\tau}}$ and $\widetilde{\theta_{B,\tau}}$. Klochkov, Härdle, and Xu (2019) show that the distribution of this test conditional on the data mimics the "true" distribution of T_1 with high probability (see Theorem 1 in the referenced script). Using (7) we can obtain the critical value through simulations. Indeed, the critical value $\mathfrak{z}_I(\alpha)$ is defined as:

$$z_I^\circ(\alpha) = z_I^\circ(\alpha; Y) = \inf\left\{z \geq 0 : \mathrm{P}(\mathcal{T}_I^\circ > z^2/2) \leq \alpha\right\}.$$

Algorithm

The algorithm for an adaptive window length selection at each point in time is now straightforward. It is based on sequential testing of the hypotheses on a nested set of intervals $\{I_k\}_{k=0,1,\ldots,K}$, where $I_0 \subset I_1 \subset \ldots \subset I_K$. Let $n_k = |I_k|$ be the number of observations in each interval. The first interval I_0 is assumed to be homogeneous with length n_0. Then, for each interval I_k, the null hypothesis of parameter homogeneity is tested against the alternative of a break point at an unknown location τ within I_k, as in (3).

Since the setup deals with nested intervals, but the existence and location of a break point are unknown, only additional points in each new interval are considered as possible break points. The candidate set for break points in each interval is defined as $J_k = I_k \backslash I_{k-1}$. Using each point $\tau \in J_k$, the left and right intervals are constructed as $A_{k,\tau} = [i_0 - n_{k+1}, \tau\,]$ and $B_{k,\tau} = (\tau, i_0]$ respectively (see Figure (3)). The test statistic is calculated similarly to equation (5) as

$$\mathcal{T}_{I_k,\tau} = L_{A_{k,\tau}}(\tilde{\theta}_{A_{k,\tau}}) + L_{B_{k,\tau}}(\tilde{\theta}_{B_{k,\tau}}) - L_{I_{k+1}}(\tilde{\theta}_{I_{k+1}}), \quad (8)$$

where $A_{k,\tau}$ and $B_{k,\tau}$ are as previously specified and we test at every point $\tau \in J_k$ for a break point. The kth interval is rejected if

$$\max_{j \in J_k} \mathcal{T}_{I_k,\tau} \leq \mathfrak{z}_{I_k}^\circ(\alpha), \quad (9)$$

and $\mathfrak{z}_{I_k}^\circ$ is generated via multiplier bootstrap as explained in the previous section. If the interval I is not rejected, i.e. there exists no break point and it is homogeneous, we continue the testing procedure by choosing a bigger interval. Otherwise, the length of the last non-rejected interval $\hat{I}$ is the interval of homogeneity and $\widehat{\theta_{\iota_0}} = \widehat{\theta_{\hat{I}}}$ is the respective adaptive estimate of $\hat{I}$.

Homogeneity testing for I_k utilizes also part of observations of I_{k+1}. Hence, the predefinition of intervals is crucial. Following that, the choice of interval lengths affects the test results, and therefore requires careful selection. Based on the initial length n_0, the interval lengths are defined by

$$n_k = \left\lceil n_0 c^k \right\rceil, \quad (10)$$

where c is a geometric multiplier, chosen slightly above 1 to ensure a monotonic increase of interval lengths, but not by a big margin. Furthermore, we select K to be the smallest integer such that the whole time series is covered under a geometrically increasing length.

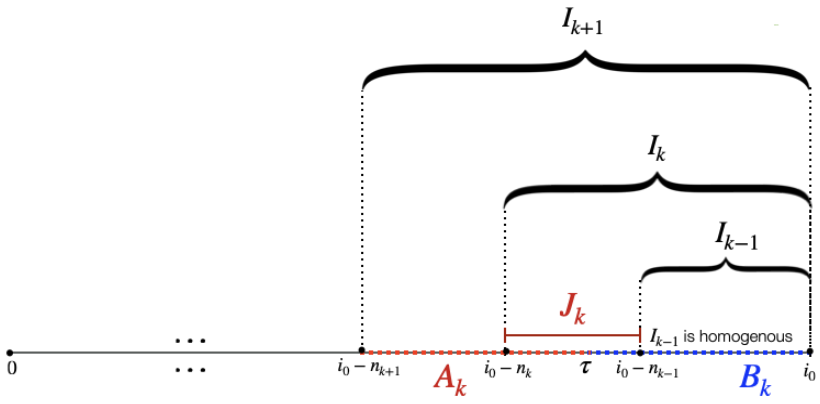


Figure 3. Iterative algorithm

Source: Authors' own research.

In summary, the LPA algorithm for adaptive choice of an interval of homogeneity and the corresponding MLE is given by the following iterative procedure:

(1) **Initialization**: Select I_0, I_1, I_2, and define $J_1 = I_1 \backslash I_0$, $(\forall \tau \in J_1)$, $A_{1,\tau} = [i_0 - n_2, \tau]$, $B_{1,\tau} = (\tau, i_0]$.

(2) **Iteration**: For each iteration, select I_{k-1}, I_k, I_{k+1}, $J_k = I_k \backslash I_{k-1}$, $(\forall \tau \in J_k)$ $A_{k,\tau} = [i_0 - nk+1, \tau]$, $B_{k,\tau} = (\tau, i_0]$.

(3) **Testing homogeneity**: Calculate test statistics in equation (8) and select critical value with multiplier bootstrap. Test hypothesis in equation (3) using equation (9).

(4) **Loop**: If I_k is accepted, take the next interval I_{k+1}. Otherwise set I_b to the latest non rejected I_k.

(5) **Adaptive estimator**: Take interval $I_b$as interval of homogeneity and $\widehat{\theta_{i_0}} = \widehat{\theta_{\hat{I}}}$ as adaptive estimate of I_b. Repeat the procedure for each point in time (different i_0).

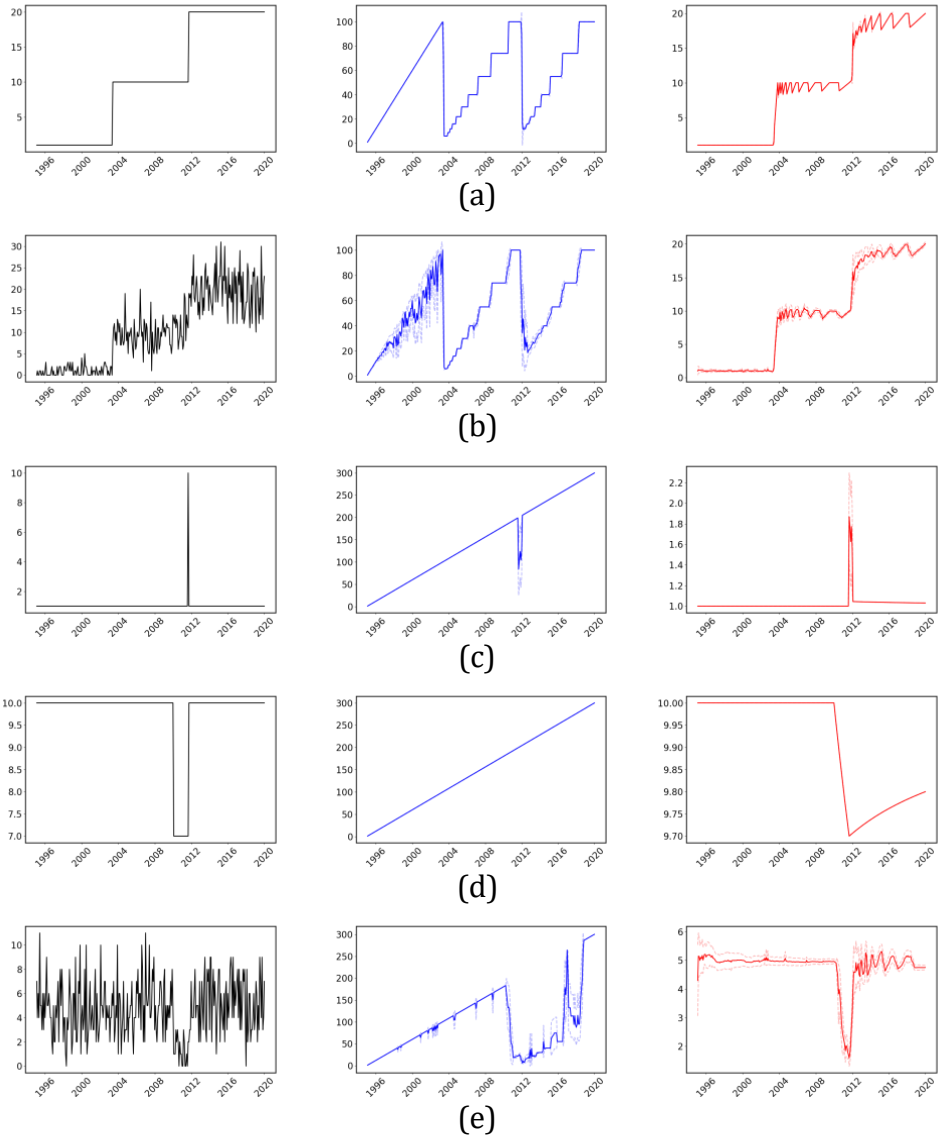
Experimental results

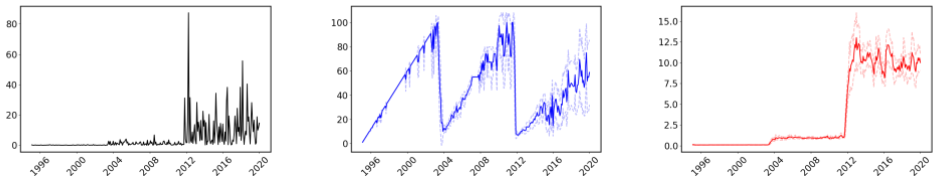
The following lines seek to answer the question whether or not we can generate better forecasts by using the described adaptive methodology, and whether or not the methodology is robust with respect to previously unseen and thus unpredictable patterns. We compare one-step-ahead and multi-period point parameter forecasts of the proposed locally adaptive procedure to a baseline of one year and three years moving averages in a pseudo-out-of-sample approach. We acknowledge that this is only a baseline evaluation approach. A more sophisticated evaluation approach would be to generate multi-period density forecasts that could be evaluated using a tailored loss-function as in Diebold, Gunther, and Tay (1998); Diebold (2015), and Gonzalez-Rivera and Sun (2014). However, as this is beyond the scope of this paper (or altogether another paper), we leave it open for future work.

Simulation study

This section evaluates the performance of the proposed technique using simulated data sets as shown in Figure 4. We create scenarios that mimic common patterns in financial data sets. The simulations focus both on short-term shocks and regime shifts. Starting with the simplest piecewise constant model, we gradually increase the complexity of the simulations by generating Poisson distributed data and finally test the robustness of the methodology by changing the underlying model to follow an exponential distribution. For each scenario, we consider a time series $(Y_t)_{t=1}^{300}$ with the following specifications:

(a) Regime shifts with piece-wise constant model: $(Y_t)_{t=1}^{100} = 1$, $(Y_t)_{t=101}^{200} = 10$ and $(Y_t)_{t=201}^{300} = 20$,
(b) Regime shifts with Poisson model: $(Y_{1t})_{t=1}^{100}$ with θ_1=1, $(Y_{2t})_{t=101}^{200}$ with θ_2=10 and $(Y_{3t})_{t=201}^{300}$ with θ_3=20,
(c) Short term shock with piece-wise constant model: $(Y_t)_{t=1}^{199} = 1$, $(Y_t)_{t=200}^{200} = 10$ and $(Y_t)_{t=201}^{300} = 1$,
(d) Structural break with piece-wise constant model: $(Y_t)_{t=1}^{180} = 10$, $(Y_t)_{t=181}^{200} = 7$ and $(Y_t)_{t=201}^{300} = 10$,
(e) Structural break with Poisson model: $(Y_{1t})_{t=1}^{180}$ with θ_1=5, $(Y_{2t})_{t=181}^{200}$ with θ_2=1 and $(Y_{3t})_{t=201}^{300}$ with θ_3=10,
(f) Regime shifts with exponential model: $(Y_{1t})_{t=1}^{100}$ with θ_1=0.1, $(Y_{2t})_{t=101}^{200}$ with θ_2=1 and $(Y_{3t})_{t=201}^{300}$ with θ_3=10.





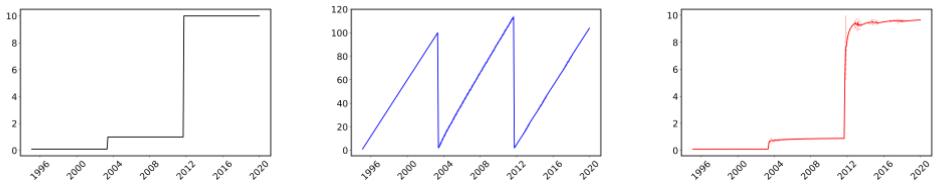
(f)

Figure 4. Simulated series (left), homogeneous windows (middle) and MLE (right). [LPA Simulations]

Source: Authors' own research.

The simulated data of scenarios (a)-(f) is shown in the time series plots on the left hand side in Figure (4). These simulations are reproducible via *quantlet.com*. Since the algorithm requires pre-selection of the intervals, we fix c in equation (10) at 1.35. Here, $c>1$ is chosen slightly greater than one, so that in the worst case one only neglects a small proportion of unknown homogeneous interval. A different selection of c will not affect the results, as long as it is not chosen to be much higher than 1. Assuming a minimal homogeneous window (n_0) of 5 months, we compute K as described in section (0.5). For example, the candidate homogeneous windows for the latest period in the time series are [6,9,12,16,22,30,40,55,74,100,135,183,247,300]. Note that since we allow the number of intervals to vary with the time point for which the homogeneous windows has to be determined, the number of intervals being tested gradually decrease as the number of historical data points in the time series decrease. The algorithm is applied to find the homogeneous window and corresponding MLE estimate for each point in time in each simulation and the results are presented in Figure (4). Also, for each scenario, we run 100 simulations to smoothen the estimated window sizes and construct confidence intervals that are depicted by dotted lines in all plots in Figure (4).

Plot (a) in Figure (4) shows that the procedure detects intervals of homogeneity correctly in all cases of this fairly simple scenario. It is robust to small and big values and gives a good approximation of the true parameters. The step-wise increase of homogeneous windows and the fluctuation of MLEs is a consequence of the limited number of possible time windows K that the algorithm tests for. If K were to be increased, for example arithmetically, which means that the algorithms tests for all possible interval sizes, increasing the interval by one at each iteration, we would expect the algorithm to generate a straight downward slopping line instead. The arithmetic increase in intervals requires a lot of tests to be performed, whereas the geometric increase with lesser number of tests already results in high accuracy. For computational ease, we illustrate this on scenario (a) in Figure (5) and avoid such an experiment for other scenarios. Similarly, we will use geometrically increasing intervals in our use-case study in the next section.



(a)

Figure 5. Simulated series (left), homogeneous windows (middle) and MLE (right) using arithmetically increasing intervals in LPA.

Source: Authors' own research.

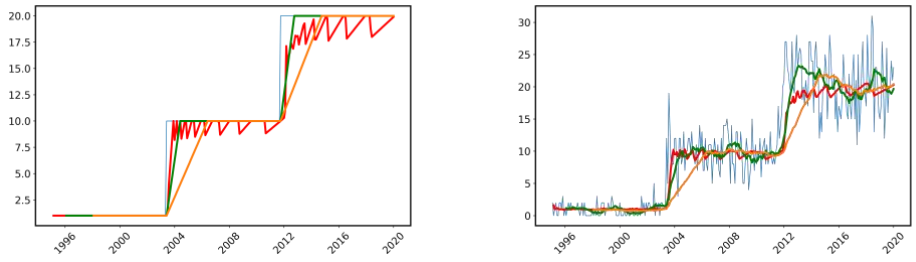
Next, a more realistic scenario with values generated from a Poisson process in (b) shows that the procedure is robust to noise even when the sample size is small. On the other hand, when n is small, the problem of multiple testing causes many false negatives. This becomes evident in the first third of (b). This can be easily dealt with by changing the number of tested intervals in the algorithm. Regardless of this issue, the MLE remains close to the true parameter value.

Furthermore, some temporary shocks and small structural breaks are mimicked in simulations (c)-(e). (c) shows that a large shock is detected accurately, with an indicator to select a smaller window size around the change in mean. The simulation in (d) shows that temporary changes in mean are not always recognised by the algorithm, but the MLE after such short-term shocks is still affected. The procedure also detects structural breaks within a Poisson framework accurately as depicted in (e), but the estimated time windows are inaccurate shortly after the break. With increasing n, the accuracy is restored and the approximated MLE remains accurate.

Finally, to check the robustness of the algorithm, values are generated from an exponential distribution. The simulation in (f) shows that the algorithm detects breakpoints correctly even with exponentially distributed data. Accordingly, the assumption of Poisson distributed values can be easily relaxed towards models from any exponential family. This robustness allows the algorithm to be used even with misspecified models.

For each simulation scenario, we also compare the results with fixed window estimates of 12 months (high variance) and 36 months (high bias) and show the results in Figure (6). All subplots show that LPA finds a balance between variance and bias by selecting a smaller or bigger window size when necessary. Plot (a) shows that while the estimates of LPA fluctuate (as described above), the shift in regime is detected earlier by LPA than by fixed window estimates. Plots (b) and (e) show that for the simulated intervals with constant mean, fixed window estimates indicate a highly fluctuating mean, while LPA recognizes intervals of time homogeneity correctly. Moreover, the impact of the short term shock disappears faster in the LPA estimates in (c) and (d). However, LPA underestimates the magnitude of shocks as visible in (d).

In conclusion, the algorithm accurately detects small shocks, regime shifts and temporary mean changes in time series without strict assumptions on the underlying model. The results could be improved further by correct selection of interval lengths and numbers of intervals being tested, as they have a significant impact on the precision and accuracy of the results.



(a) (b)

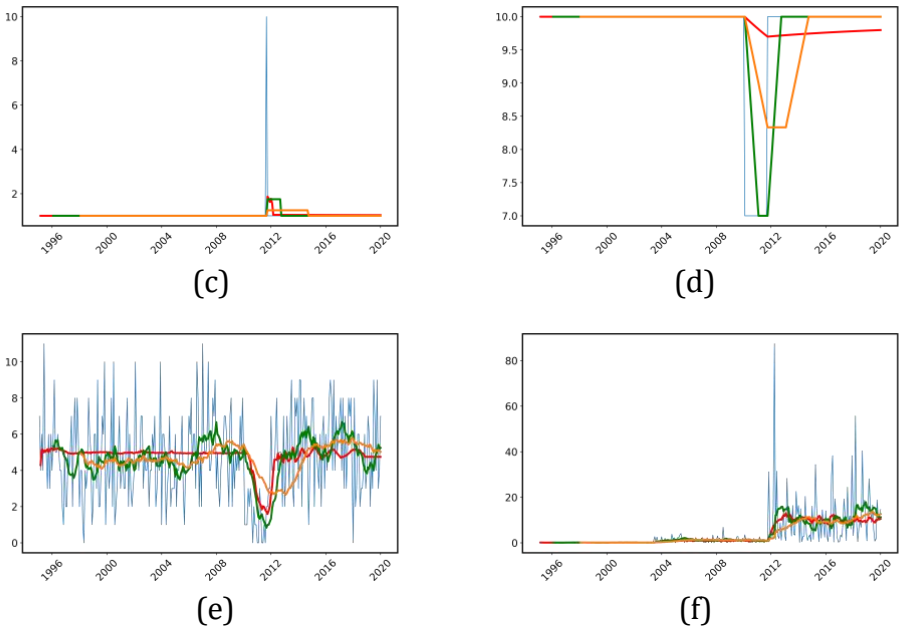


Figure 6. Time series of simulated data and one step ahead prediction with estimates from LPA, 1 year fixed window and 3 year fixed window

Source: Authors' own research.

Use case study

In this section, we apply LPA to a working example of time series that are relevant to businesses in the financial industry. We use a data set of mergers & acquisitions per month that we acquired from the database Refinitiv Eikon Deals (restricted access through subscription). We consider different industries according to the Thomson Reuters Business Classification scheme in Germany. The data set consists of a total of 9,969 observations in ten industries in the US and Germany. We put the procedure into action on the following three industries in Germany:

(1) Financials
(2) Telecommunication
(3) Energy

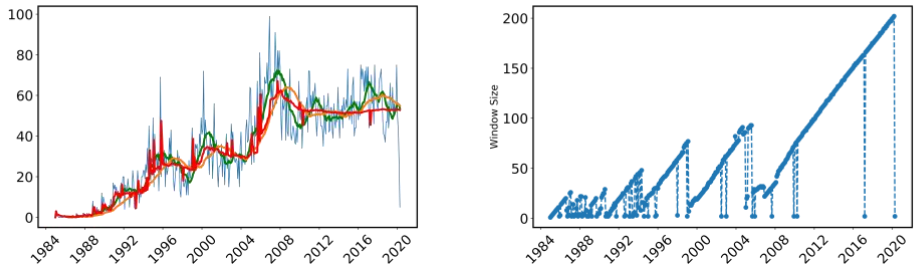
Due to inconsistent recordings of transactions, we considered only 424 observations from 01-1984 to 04-2020 in Germany for each industry. Following previous literature, we consider different industries separately. Without insider knowledge, the frequency of M&A events could be assumed to be random, which is based on the intuition that we usually learn about mergers only after they are announced. We also assume that at each point, there exist a time window where a somewhat stable structure can be found in the short run, but can be interrupted by exogenous factors that change industry dynamics. Given these characteristics, we model the number of M&A with a Poisson process and apply LPA to forecast the next period given the MLE of the last observed homogeneous time window. We do not pre-process the data and let the procedure account for structural breaks and non-stationarity.

Similar to the simulation study in previous section, we fix n_0 at 5 months, and c at 1.35. To re-iterate, this seemingly arbitrary choice ensures that we neglect only few unknown homogeneous intervals (if any), while keeping the number of performed tests minimal. Next, we run 100 simulations to smoothen the estimated window sizes. We keep n_0 and c for the

interval selection through equation (10) constant as an aim of the proposed algorithm is to evaluate unseen data even if no knowledge about it is available. Some example cases are illustrated to verify the robustness of the algorithm with respect to different types of time series in Figure (7). The plot on the right hand side shows the homogeneous windows for each point in time. The corresponding LPA estimate, 12 month estimate, and 36 month estimate as one period ahead forecast are shown in the left plot.

The first plot in Figure (7) shows the German financial industry, where an upward trend in the number of mergers can be observed over time. The plot shows that LPA estimates closely mimic this trend, and is much more responsive to shocks as in the mid-90s. Moreover, contrary to the green 1-year moving average curve, which shows high variance in the MLE estimate, LPA recognizes the regions within time series where the average number of M&A is approximately constant, such as from 2010 to 2020. The LPA estimates on the left suggest that the selected window size differs over time (sometimes up to 200, sometimes only a few observations) and using a fixed time window for generating forecasts is not recommended according to the technique. Comparing the window plot of the same industry with the simulation plots (a) and (b) from the previous section, we identify six regimes in the German financial industry (1984-1988, 19881990, 1990-1994, 1994-1999, 1999-2006 and 2006-2020). These regimes could be associated with external events, such as: a merger wave in 1984, the effects of globalization in the late 1980s, the German bank merger wave in the 1990s, market liberalization policies in the mid-90s, and the technological innovations in the late 20th century.

The second part of Figure (7) shows the German telecommunication industry, where the number of M&As per month seems stable and stationary, except for a short merger wave in the mid-90s due to the liberalization of the market and another short wave around 2001 (possibly related to the dotcom-bubble). LPA suggests a stable time series as the algorithm recommends to select large windows, and sometimes the whole time series. Only during the interval where a shock can be seen on the left, LPA's choice of homogeneous window on the right restricts the time window to a few years (for example between 1994 and 2001). We see that LPA can handle time-varying parameters while facing a trade-off between parameter variability and modelling bias. A similar conclusion can be drawn from the energy markets plots in (3) of Figure (7).



(1)

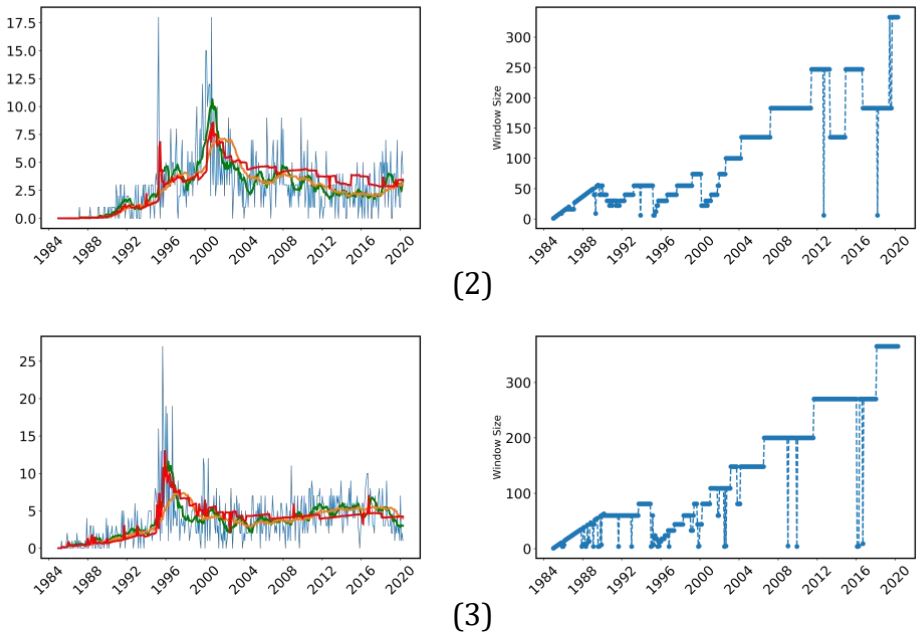
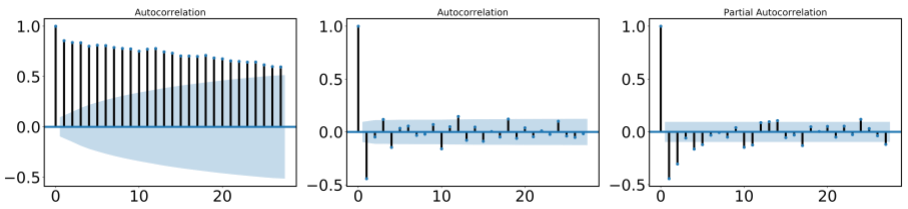


Figure 7. Time series of original data and one step ahead prediction with estimates from LPA, 1 year fixed window and 3 year fixed window (left) Homogeneous windows (right) for (1) Financials, (2) Telecommunication and (3) Energy

Source: Authors' own research.

We further extend our analysis by comparing the predictions from our LPA approach with a baseline ARIMA model. We use the *auto.arima* function in R that uses the algorithm given in Hyndman and Khandakar (2008) for finding the best fit for our data. The best model for all three datasets is summarized in table (1). As expected, all three series show non-stationarity and require differencing. We verify this by examining the auto-correlation function plots of the series as shown in Figure (8). The auto-correlation lags are positive for many lags, this indicates that the series needs differencing. The second and third panel of the same plot show that the time series becomes stationarity after taking first differences. Furthermore, the models suggest no autoregressive term, and only one or two moving average terms. While these models have led to a small error for the forecast period of last three months, they can not be relied on when there is an abrupt change in the time series, e.g. as in 1996. We repeat the procedure the analysis for the period from 01-1984 to 09-1996 and use 10-1996 to 12-1996 as forecast period.

The results reveal that the best fit ARIMA model for this time period changes for all three data sets (ARIMA(2,1,2), ARIMA(1,1,2), and ARIMA(0,1,2) respectively), suggesting different autoregressive and moving average lags. This implies that a single parametric model is not a good fit for the entire time series.



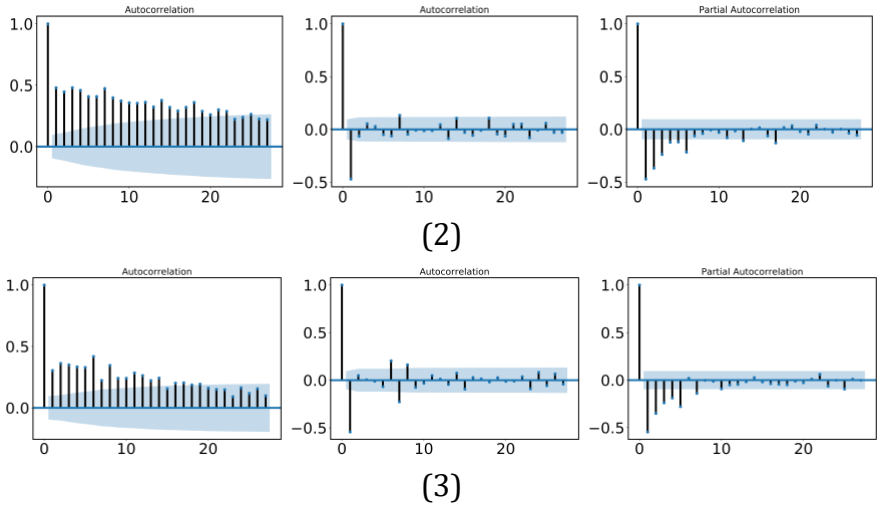


Figure 8. Autocorrelation plot for the original data (left), autocorrelation plot for the first differenced data (middle) and partial autocorrelation plot (right) for (1) Financials, (2) Telecommunication and (3) Energy

Source: Authors' own research.

Next, we forecast the number of M&As for 02-2020 to 04-2020 using the LPA estimate and the fixed 1-year and 3-year window estimates of 01-2020. The estimates and mean are summarized in table (1). Apart from the selection of time-varying window lengths, the results suggest that LPA can also provide guidance on the a-priori selection of fixed windows. Looking at the distribution plot of window lengths (as a proportion of data points in the time series before the time of evaluation) in Figure (9), and selecting the window with highest frequency (w) as the fixed time window for each time series, we also make a forecast reported in the same table.

In assessing the quality of our forecasts, we rely on two widely used measures: Mean Squared Error (MSE) and Mean Absolute Error (MAE). These metrics are commonly employed in evaluating short-term forecast accuracy across various fields due to their intuitive interpretation and robustness. While methods such as the Diebold-Mariano test and White's test are often utilized for comparing forecast accuracy between different models, they may not be suitable for our analysis due to certain limitations or assumptions. For instance, the Diebold-Mariano test requires normally distributed forecast errors and stationary time series, which may not always hold true in our context. Similarly, White's test assumes homoscedasticity of errors, condition that may not be met in our dataset. Therefore, we refrain from employing these tests and instead focus on MSE and MAE as reliable indicators of forecast performance, providing valuable insights into the accuracy and reliability of our forecasting methods. These measures for all the above methods are presented in table (2).

For completion, we also compare the predictions from our LPA approach with a BCP method using a Poisson likelihood. For the BCP method, we identify the last change point as the breakpoint and use the mean value of the time series from this point onwards as the forecast for subsequent time points as reported in table (1). We apply this approach across all three industries and report the results in the same table. Moreover, we calculate MSE for the BCP method to provide a comprehensive comparison, see table (2). The inclusion of the

BCP method allows us to evaluate the robustness of our LPA approach against a standard benchmark, especially in scenarios involving abrupt changes in the time series.

Table 1. Forecast results for 02-2020 to 04-2020 based on the adaptively selected MLE, fixed windows, BCP, ARIMA, and most recurring window proportionally (w) in the distribution plots in Figure (9)

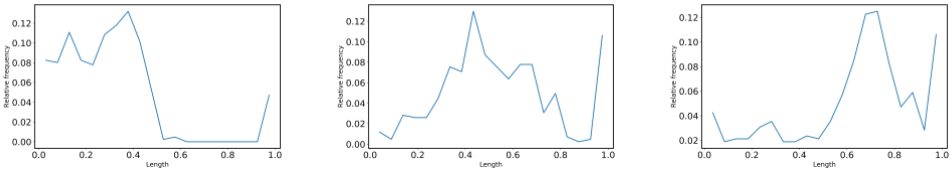
	Financials	**Telecom.**	**Energy**
LPA estimate	52.99	3.43	4.23
12 month MA estimate	55.58	2.42	3.08
36 month MA estimate	55.50	2.94	4.17
BCP estimate	75.0	1.0	1.0
Most recurring window(w)	160	184	309
w month MA estimate	55.25	2.90	4.62
w month MSE	1139.23	5.06	11.03
ARIMA best fit	(0,1,2)	(0,1,1)	(0,1,1)
ARIMA MSE	301.15	0.79	0.23

Source: Authors' own research.

Table 2. Forecast accuracy in terms of MAE and MSE calculated for 02-2020 to 04-2020 based on the adaptively selected MLE, fixed windows, and BCP

	Financials	**Telecom.**	**Energy**
LPA MAE	27.65	1.56	2.90
1-year fixed MAE	30.25	2.58	1.75
3-year fixed MAE	30.17	2.06	2.83
BCP MAE	49.67	4.0	0.33
LPA MSE	1008.84	3.11	8.61
1-year fixed MSE	1159.28	7.34	3.29
3-year fixed MSE	1154.25	4.89	8.25
BCP MSE	2711.0	16.67	0.33

Source: Authors' own research.



(1) (2) (3)

Figure 9. Distribution of estimated interval length as a proportion of data points for (1) Financials, (2) Telecommunication and (3) Energy

Source: Authors' own research.

Table (2) shows that homogeneous time window recommended by LPA for the financial industry was roughly 160 months. Choosing this time window resulted in a lower MSE than using 12 and 36 months fixed window estimates. For the energy sector, however, LPA recommends the selection of a very long window, for which the estimate deviates

significantly from the true value, but still captures the stationary aspect of the time series. Table (2) shows that LPA produces the smallest MSE and MAE in the financials and telecommunication industry. However, in the energy sector the 1 year fixed window and BCP outperforms LPA, perhaps due to high fluctuations in this industry. As a whole, the results show that cluster patterns and level shifts are accurately detected, and that a plausible fit is achieved under the assumption that M&A follow a Poisson distribution with time-varying parameter.

The analysis shows that the proposed technique efficiently handles the complex properties of M&A time series and helps us generate better forecasts. Moreover, the empirical analysis reveals that the patterns in the various German industries differ strongly. Therefore, the usage of an automated and data driven procedure lets us obtain accurate forecasts with minimal effort and can be adapted to other industries or even other business problems. It can be used as a baseline for forecasting or simulation approaches and can be integrated into large scale planning systems. The MLE series could be forecasted using e.g. autoregressive models that generate different scenarios. Business experts could then make a judgement about the likeliness of these scenarios. Moreover, the availability of a time-varying parameter allows analysts to use more sophisticated approaches, e.g. modeling the distributional properties of various time-series that are assumed to follow a Poisson distribution, such as the examples mentioned in the introduction (call centers, sales, supply chains). Businesses are often not interested in exact values, but in having an overview over the range of expected figures, and forecasts of low quality can lead to costly miscalculations. The proposed procedure helps increase forecast quality and availability and thus generate business value if employed.

Conclusion

A new algorithm for automatically finding homogeneous time intervals in a non stationary count time series data is proposed. Trends, cyclical components, and even black swan events can be tackled with this technique, without need of manual supervision or model assumptions. The algorithm is based on a fruitful combination of a local parametric model (here a Poisson process) and MBS. We conduct a simulation study that indicates that the procedure is indeed robust, even if the input data is not Poisson distributed. These results are then used to evaluate a data set of M&A industry-based in Germany. The results show that this non-parametric model provides accurate forecasts and also helps in finding an appropriate window size for other models. In conclusion, we provide an easy-to use solution for analyzing large data collections automatically that addresses the most common problems in time series modeling and can be adapted to diverse applications, including call center arrivals, sales forecasting or supply chain decisions.

This work is subject to several limitations that are caused by additional complexity that exceeds the scope of this work. We evaluated only three German industries, chosen randomly. However, data for all industries in the US and Germany is available and could be used. The simulations showed that the proposed method had difficulties with detecting small changes in distribution. We also saw that the method is not in all cases better than choosing a simple approach using moving averages. An evaluation based on all 20 data sets could give further insights on the robustness of the method. Furthermore, the evaluation using a pseudo-out-ofsample point forecast is fairly limited and we justify the superiority of our approach based on the simple measure of a mean-squared-error. Diebold (2015) outlines

that this approach provides no insurance against over-fitting and is computationally costly. Another questionable point is whether the occurrence of M&A per month strictly follows a time-varying Poisson process, as assumed in this paper. Fortunately, the simulation study shows that this assumption need not necessarily be fulfilled though.

Answering these questions using a sophisticated evaluation approach through generating (multi-period) density forecasts would give more insights and could be an interesting area for future research. Future research could also evaluate the fit to a Poisson distribution with timevarying MLE, but with respect to different values of n0 and c. Alternatively, other methods for determining the optimal number of tests to be performed to find balance between computational cost and accuracy can be developed. Finally, a possible extension could include a more general version of the current test statistic that looks into both directions.

Acknowledgment

Support from the German Research Foundation [IRTG 1792] greatly acknowledged; the European Union's Horizon 2020 "FIN-TECH" Project [825215, Topic ICT-35-2018, Type of actions: CSA]; the European Cooperation in Science & Technology COST Action [CA19130]; the Czech Science Foundation [19-28231X, CAS: XDA 23020303]; and the Yushan Scholar Program of Taiwan.

References

Ahern, K. R., and Harford, J., (2014), "The Importance of Industry Links in Merger Waves," The Journal of Finance, Vol. 69, No. 2, pp. 527–576.

Akkus, O., Cookson, J. A., and Hortac¸su, A., (2015), "The Determinants of Bank Mergers: A Revealed Preference Analysis," Management Science, Vol. 62, No. 8, pp. 2241–2258.

Aminikhanghahi, S., and Cook, D. J., (2017), "A Survey of Methods for Time Series Change Point Detection," Knowledge and Information Systems, Vol. 51, No. 2, pp. 339–367.

Andrews, D. W. K., and Ploberger, W., (1994), "Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative," Econometrica, Vol. 62, No. 6, pp. 1383–1414.

Arlot, S., Blanchard, G., and Roquain, E., (2010), "Some nonasymptotic results on resampling in high dimension, I: Confidence regions," Annals of Statistics, Vol. 38, No. 1, pp. 51–82.

Beran, R., (1986), "Discussion: Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis," Annals of Statistics, Vol. 14, No. 4, pp. 1295–1298.

Bianchi, D., and Chiarella, C., (2018), "An Anatomy of Industry Merger Waves," Journal of Financial Econometrics, Vol. 17, No. 2, pp. 153–179.

Bücher, A., and Dette, H., (2013), "Multiplier bootstrap of tail copulas with applications," Bernoulli, Vol. 19, No. 5A, pp. 1655–1687.

Chatterjee, S., and Bose, A., (2005), "Generalized bootstrap for estimating equations," Annals of Statistics, Vol. 33, No. 1, pp. 414–436.

Chen, J., and Gupta, A. K., (1999), "Change point analysis of a Gaussian model," Statistical Papers, Vol. 40, No. 3, pp. 323–333.

Chen, J., and Gupta, A. K., (2011), Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance, Springer Science & Business Media.

Chernozhukov, V., Chetverikov, D., and Kato, K., (2013), "Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors," Annals of Statistics, Vol. 41, No. 6, pp. 2786–2819.

Diebold, F. X., (2015), "Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold–Mariano Tests," Journal of Business & Economic Statistics, Vol. 33, No. 1, pp. 1–1.

Diebold, F. X., Gunther, T. A., and Tay, A. S., (1998), "Evaluating Density Forecasts with Applications to Financial Risk Management," International Economic Review, Vol. 39, No. 4, pp. 863–883.

Eckley, I. A., Fearnhead, P., and Killick, R., (2011), "Analysis of changepoint models," August.

Giraitis, L., Kapetanios, G., and Price, S., (2013), "Adaptive forecasting in the presence of recent and ongoing structural change," Journal of Econometrics, Vol. 177, No. 2, pp. 153–170.

Gonzalez-Rivera, G., and Sun, Y., (2014), Density Forecast Evaluation in Unstable Environments. Technical Report 201428. University of California at Riverside, Department of Economics.

Haccou, P., Meelis, E., and van de Geer, S., (1987), "The likelihood ratio test for the change point problem for exponentially distributed random variables," Stochastic Processes and their Applications, Vol. 27, pp. 121–139.

Härdle, W. K., and Mammen, E., (1993), "Comparing Nonparametric Versus Parametric Regression Fits," Annals of Statistics, Vol. 21, No. 4, pp. 1926–1947.

Härdle, W. K., Hautsch, N., and Mihoci, A., (2015), "Local Adaptive Multiplicative Error Models for High-Frequency Forecasts," Journal of Applied Econometrics, Vol. 30, No. 4, pp. 529–550.

Harford, J., (2005), "What drives merger waves?" Journal of Financial Economics, Vol. 77, No. 3, pp. 529–560.

Hinkley, D. V., and Hinkley, E. A., (1970), "Inference About the Change-Point in a Sequence of Binomial Variables," Biometrika, Vol. 57, No. 3, pp. 477–488.

Hsu, D. A., (1979), "Detecting Shifts of Parameter in Gamma Sequences with Applications to Stock Price and Air Traffic Flow Analysis," Journal of the American Statistical Association, Vol. 74, No. 365, pp. 31–40.

Hyndman, R. J, and Khandakar, Y., (2008), "Automatic Time Series Forecasting: The forecast Package for R," Journal of Statistical Software, Vol. 27, No. 3, pp. 1–22.

Inoue, A., Jin, L., and Rossi, B., (2017), "Rolling window selection for out-of-sample forecasting with time-varying parameters," Journal of Econometrics, Vol. 196, No. 1, pp. 55–67.

Klochkov, Y., H¨ardle, W. K., and Xu, X., (2019), "Localizing Multivariate CAViaR," IRTG1792 Discussion paper 48.

Kolsarici, C., and Vakratsas, D., (2015), "Correcting for Misspecification in Parameter Dynamics to Improve Forecast Accuracy with Adaptively Estimated Models," Management Science, Vol. 61, No. 10, pp. 2495–2513.

Kutoyants, Y. A., and Spokoiny, V., (1999), "Optimal choice of observation window for Poisson observations," Statistics & Probability Letters, Vol. 44, No. 3, pp. 291–298.

Maksimovic, V., Phillips, G., and Yang, L., (2013), "Private and Public Merger Waves," The Journal of Finance, Vol. 68, No. 5, pp. 2177–2217.

Mammen, E., (1993), "Bootstrap and Wild Bootstrap for High Dimensional Linear Models," Annals of Statistics, Vol. 21, No. 1, pp. 255–285.

Martynova, M., and Renneboog, L., (2005), A Century of Corporate Takeovers: What Have We Learned and Where Do We Stand? (previous title: The History of M&A Activity Around the World: A Survey of Literature). SSRN Scholarly Paper ID 820984. Rochester, NY: Social Science Research Network.

Mercurio, D., and Spokoiny, V., (2004), "Statistical inference for time-inhomogeneous volatility models," Annals of Statistics, Vol. 32, No. 2, pp. 577–602.

Oreshkin, B. N., Réegnard, N., and L'Ecuyer, P., (2016), "Rate-Based Daily Arrival Process Models with Application to Call Centers," Operations Research, Vol. 64, No. 2, pp. 510–527.

Pesaran, M. H., Pick, A., and Pranovich, M., (2013), "Optimal forecasts in the presence of structural breaks," Journal of Econometrics, Vol. 177, No. 2, pp. 134–152.

Spokoiny, V. G., (1998), "Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice," The Annals of Statistics, Vol. 26, No. 4, pp. 1356–1378.

Spokoiny, V., (2009), "Multiscale local change point detection with applications to value-at-risk," The Annals of Statistics, Vol. 37, No. 3, pp. 1405–1436.

Spokoiny, V., and Zhilova, M., (2015), "Bootstrap confidence sets under model misspecification," Annals of Statistics, Vol. 43, No. 6, pp. 2653–2675.

Taylor, J. W., (2007), "A Comparison of Univariate Time Series Methods for Forecasting Intraday Arrivals at a Call Center," Management Science, Vol. 54, No. 2

Taylor, J. W., (2011), "Density Forecasting of Intraday Call Center Arrivals Using Models Based on Exponential Smoothing," Management Science, Vol. 58, No. 3, pp. 534–549.

Very, P., Metais, E., Lo, S., and Hourquet, P. G., (2012), "Can We Predict M&A Activity?" In Advances in Mergers and Acquisitions, edited by Sydney Finkelstein and Cary L. Cooper, Vol. 11 of Advances in Mergers & Acquisitions, pp. 1–32, Emerald Group Publishing Limited.

Wu, C. F. J., (1986), "Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis," Annals of Statistics, Vol. 14, No. 4, pp. 1261–1295.

Yelland, P. M., Kim, S., and Stratulate, R., (2010), "A Bayesian Model for Sales Forecasting at Sun Microsystems," Interfaces, Vol. 40, No. 2, pp. 118–129.