COVID-19 contagion and digital finance

Arianna Agosto[1] · Paolo Giudici[1]

Abstract

Digital finance is going to be heavily affected by the COVID-19 outbreak. We present a statistical model which can be employed to understand the contagion dynamics of the COVID-19, so that its impact on finance can possibly be anticipated, and digitally monitored. The model is a Poisson autoregression of the daily new observed cases, and considers both short-term and long-term dependence in the infections counts. Model results are presented for the observed time series of China, the first affected country, but can be easily reproduced for all countries.

Keywords Contagion monitoring · Poisson autoregressive models · Financial crisis

JEL classification C11 · C15 · C51 · C52 · C55 · C58 · G01 · G12

1 Background

It is well known that financial markets and institutions are heavily exposed to stressful events. Digital finance can help in making available in real time, and possibly on a smartphone, the results obtained from a massive data analysis aimed at establishing the systemic consequences of a stressful event. This is clearly exemplified in the recent work of Yu et al. (2019), who show how artificial intelligence, econometrics, statistics and fast parallel computing can join their efforts to build a Financial Risk Meter (FRM, see Mihoci et al. 2020), that provides an easy-to-read and interpretable monitor of the impact of stressful events in finance. Such a monitor can be applied not only by policy makers, to take informed decisions, but also by citizens, to increase the knowledge and understanding of stress events, and of their impact, so to improve awareness and, eventually, acceptance of measures aimed at contrasting the source of such stress.

As we are writing, the world is impacted by the COVID-19 outbreak stress, which is having a huge impact, on markets and institutions. Digital finance is

✉ Paolo Giudici
giudici@unipv.it

[1] University of Pavia, Pavia, Italy

already incorporating this effect, and giving relevant information to decision makers and to the public. This activity can be exemplified by two analyses that are constantly updated on-line, that were kindly shared with us by the original sources.

The first example concerns the impact of COVID-19 on credit risk, which is being estimated by Pediroda (2020). We report the related results, to date, in Fig. 1. The figure shows the results of a simulation study that assumes two extreme scenarios for Italian SME companies: (a) a 4% yearly decrease in turnover; (b) a more extreme (indeed realistic) scenario, which assumes that turnover decreases by 10%. For both scenarios, the change in the rating classification of Italian companies is reported.

Figure 1 clearly shows that the largest impact of the COVID-19 outbreak will be on the intermediate risk classes (B, BB, BBB, which correspond to about 65% of Italian SMEs), for which the default probability could go from 0.98 to 2.14% (under scenario a) and possibly from 0.98 to 3.29% (under scenario b). We remark that the results in Fig. 1 are constantly updated, and will be publicly reported.

The second example concerns the impact of COVID-19 on market risk, which is being estimated by the cited FRM (see Mihoci et al. 2020). We report the results of FRM application to European financial markets, in Fig. 2.

Figure 2 clearly shows the impact of the COVID-19 on the European financial markets, in two subsequent phases. The first one starts on February 21st, when the news about the outbreak of local contagion in Italy becomes of public domain. The second starts on February 27th, when many positive contagion cases are also reported for France, Germany and the UK. In both cases, there is a sudden increase in the systematic risk values, both in their main body (the boxplots, which also show an increased variability between the constituent markets) and in their maximum values (the red line on top). Again, we remark that the results in Fig. 2 are constantly updated, and publicly reported, at: http://frm.wiwi.hu-berlin.de.
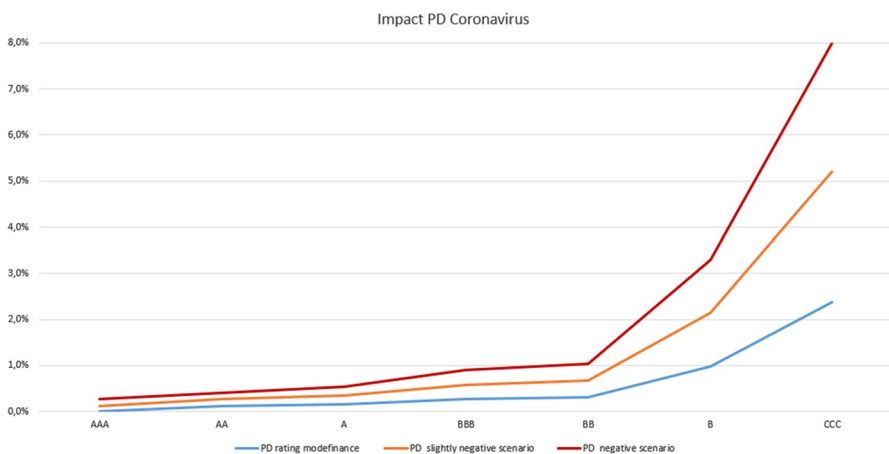
Fig. 1 Estimated impact of COVID-19 on Italian SME ratings. Source: Valentino Pediroda, University of Trieste
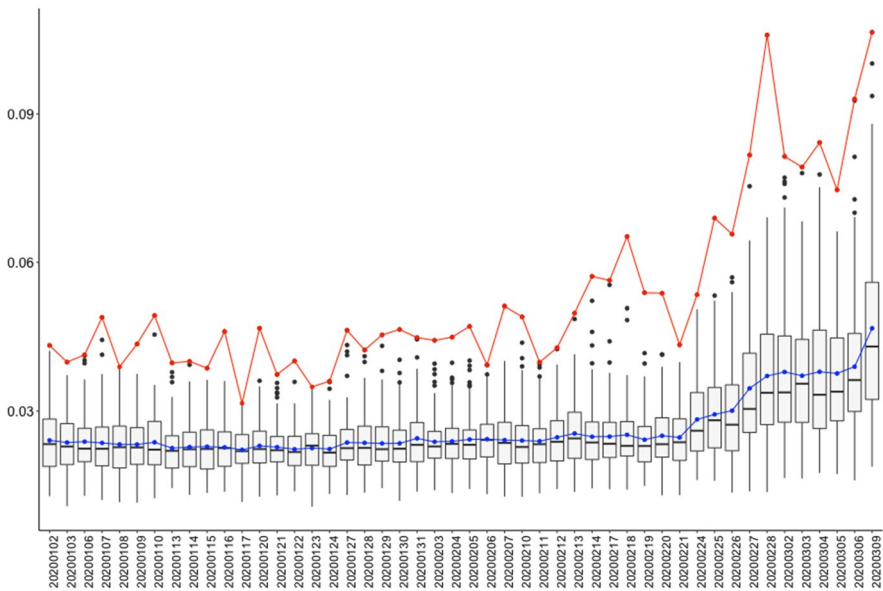
Fig. 2 Estimated impact of COVID-19 on European financial markets. Source: http://frm.wiwi.hu-berlin.de

From the previous considerations, it seems that the impact of the COVID-19 outbreak on companies, financial institutions and markets is high. On the other hand, it is clear that this impact depends on what is the perceived effect of contagion, especially at the country level, as this is the level at which most policy-making decisions are taken. We believe that digital finance can, once more, help, by attempting to provide predictive models for COVID-19 diffusion for each country which can be made accessible on-line, to better inform the decisions of policy makers and to improve the awareness and the willingness to accept restrictive measures, also on the economic side. In addition, the same contagion diffusion measurements could be implemented in risk monitors, as the ones shown before, to improve predictions.

There is, currently, lack of predictive models, particularly in terms of the econometric standards to which digital finance is used. This paper proposes to fill this gap, with a predictive model which is taken from the econometric literature on default contagion, and adapted to the epidemic context, a model that is also easy to explain and interpret and, therefore, to implement in a digital finance interface. The next section presents the proposed model. Section 3, instead, presents the results from the application of the model to the available time series for China. The data, available upon request to the Authors, allow the model to be reproduced and extended to other time periods and countries.

2 Methods

In this section, we aim to build a data-driven model which can provide support to policy makers engaged in contrasting the spread of COVID-19. To this aim, we propose a statistical framework to model the contagion dynamics, so that preventive measures (such as mobility restrictions) can be applied and/or relaxed.

To be built, the model requires, for each country (or region), the daily count of new infections. In the study of epidemics, it is usually assumed that infection counts follow an exponential growth, driven by the reproduction number R_0 (see, e.g., Biggerstaff et al. 2014). The latter can be estimated by the ratio between the new cases arising in consecutive days: a short-term dependence. This procedure, however, may not be adequate, as the incubation time is quite variable among individuals, and data occurrence and measurement is not uniform across different countries (and, sometimes, along time): these aspects induce a long-term dependence.

From the previous considerations, it follows that it would be ideal to model newly infected counts as a function of both a short-term and a long-term component. A model of this kind has been recently proposed by Agosto et al. (2016), in the context of financial contagion. We propose to adapt this model to the COVID-19 contagion.

Formally, resorting to the log-linear version of Poisson autoregression, introduced by Fokianos and Tjøstheim (2011), we assume that the statistical distribution of new cases at time (day) t, conditional on the information up to $t-1$, is Poisson, with a log-linear autoregressive intensity, as follows:

$$Y_t|\mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t)$$
$$\log(\lambda_t) = \omega + \alpha \log(1 + y_{t-1}) + \beta \log(\lambda_{t-1}),$$

where $y \in \mathbb{N}$, $\omega \in \mathbb{R}$, $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$. Note that the inclusion of $\log(1 + y_{t-1})$, rather than $\log(y_{t-1})$, allows to deal with zero values.

In the model, ω is the intercept term, whereas α and β express the dependence of the expected number of new infections, λ_t, on the past counts of new infections. Specifically, the α component represents the short-term dependence on the previous time point. The β part expresses instead the long-term dependence on all past values of the observed process, and can, thus, be interpreted as a trend component. Its inclusion is analogous to moving from an ARCH (Engle 1982) to a GARCH (Engle and Bollerslev 1986) model in Gaussian processes, and allows to capture long memory effects. The advantage of a log-linear intensity specification, rather than the linear one known as integer-valued GARCH (see, e.g., Ferland et al. 2006), is that it allows for negative dependence. From an inferential viewpoint, Fokianos and Tjøstheim (2011) show that the model can be estimated by a maximum likelihood method.

3 Results

We apply the model to the available data for China, the country first affected by COVID-19. The data cover the period from January 20 to March 31, 2020. The data source is the daily World Health Organisation reports (WHO, see World Health

Organisation 2020), from which we have extracted the "Total confirmed new cases". Figure 3 presents the observed evolution of the daily new cases of infection for China (starting from January, 20th).

We remark that the data source documents a change in the definition of new cases, in China, starting with the report of February 14th. On February 13th, besides laboratory-confirmed cases, also clinically suspected cases are reported, and also retrospectively, creating a data spike of more than 15,000 cases. This additional counting, and consequent data inflation, has continued until February 18th. The WHO report of February 20th states that, from February 19th, only laboratory-confirmed cases are included, exactly as before February 13th. We decided to consider only laboratory-confirmed cases, in line with World Health Organisation (2020).

Figure 3 reports the observed dynamic of the counts of daily new infected cases in China, in the considered period.

Figure 3 shows that COVID-19 contagion in China has completed a full cycle, with an upward trend, a peak, and a downward trend. The application of our model can better qualify these conclusions. The estimated model parameters for China are shown in Table 1.

Table 1 shows that all estimated autoregressive coefficients are significant, confirming the presence of both a short-term and a long-term dependence. From an interpretational viewpoint, the estimate of α shows that, if the expectation of new cases for yesterday was close to 0, 100 new cases observed yesterday generate about 40 new expected cases today. According to the value estimated for β, an expectation of 100 new cases for yesterday generates instead about 2 new expected cases today, if no cases were observed yesterday. These results show that, at the writing date, the short-term component is driving the infection count process in China, which is well beyond the peak. Additionally, we remark that the goodness of fit of the model
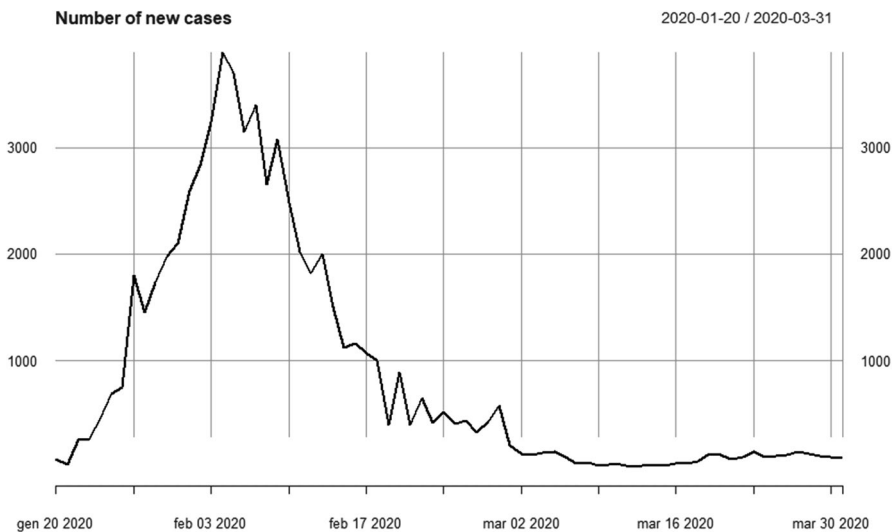
Fig. 3 Evolution of the observed infection counts in China

Table 1 Model estimates for China, with standard errors and p values of a two-sided t test

Parameter	Estimate	Std error (p value)
ω	0.337	0.247 (0.177)
α	0.823	0.069 (0.000)
β	0.133	0.062 (0.036)

is quite high, as the root mean squared error (RMSE) is equal to 278.55, against an overall mean of 868.11.

To gauge the impact of the growth in contagion counts on digital finance, we measure its direct impact on financial markets. To this aim, we have downloaded from Yahoo! Finance (https://finance.yahoo.com/) the time series of the Shangai Stock Exchange composite index (SSE), from February 20th until March 31st, in line with our analysis of contagion counts. Let p_t, $t = 1, \ldots, T$ indicate the obtained series, with t denoting a day in which the market is open.

We have transformed the index time series into a return time series, by:

$$r_t = \log \left(\frac{p_t}{p_{t-1}} \right).$$

We have then calculated, for the same days, the logarithmic variation in the daily number of contagion counts, as:

$$v_t = \log \left(\frac{y_t}{y_{t-1}} \right).$$

We remark that daily contagion counts are available also for the days in which the markets are closed. However, we assume that the signal that matters for market is the variation in the level of contagion over two consecutive open market days. This assumption is in line with the underlying rationale of epidemiologic models, whose core is the reproductive number of the virus (R_0), usually calculated from the ratio $\frac{y_t}{y_{t-m}}$, with an m that can vary according to the random variation of the incubation time. In this sense, the markets do not have special reasons to consider $m = 1$ rather than an m that is equal to the number of consecutive days of open markets ($m = 3$ during weekends, for example).

Figure 4 reports the observed time series of the SSE index returns r_t, along with the corresponding observed series of the contagion counts variation in China, v_t, for the considered period. For the sake of clarity, we reported both series with their calendar interruptions, due to weekends or vacation periods.

Figure 4 shows that the two curves are initially positively correlated: before the beginning of February the number of infected cases (as can be seen in Fig. 3) is still low, and the market does not seem to react. In the first week of February, instead, after the Chinese New Year's celebrations, contagion tops up and the SSE index plummets. Negative correlations, although weaker, are observed also during the following weeks. Indeed, the overall Pearson correlation coefficient between the two series is equal to -0.33 and significant (p value $= 0.025$).
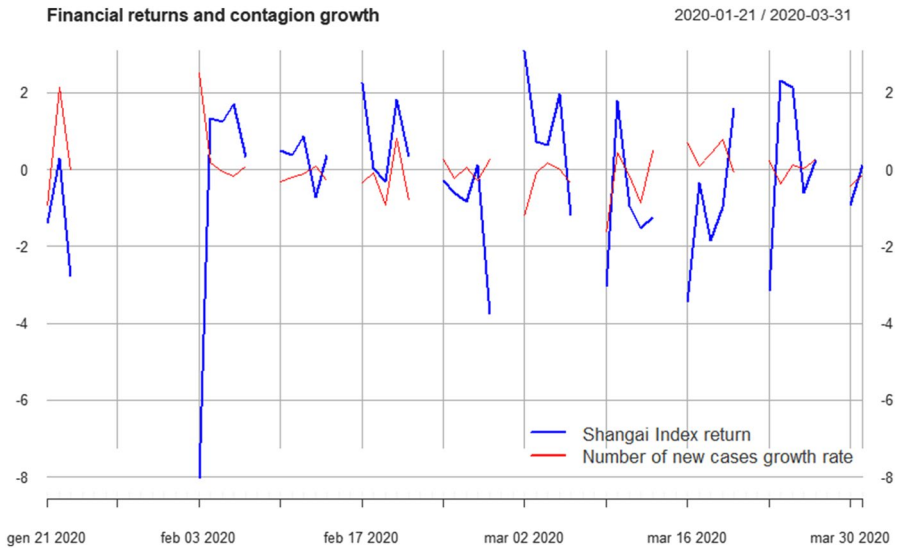
Fig. 4 Comparison between SSE returns and contagion counts variations (the SSE returns were multiplied by 100 to ensure proper comparison of the series). The series are interrupted by market closure days

We remark the count variations in Fig. 4 are based on the observed counts, which may be affected by daily random variations. A good statistical model for disease counts should be able to capture the core part of the counts variation, smoothing away the random component. To verify this statement, we also consider the relationship between the financial market returns and the fitted values of the proposed Poisson autoregressive model. Such values are plotted in Fig. 5, along with the SSE index returns, similarly as before.

Compared with Figs. 4, 5 shows a more evident correlation. Model smoothing helps to increase the overall correlation, which indeed increases up to −0.48 and is, again, significant (p value = 0.001).

4 Conclusions

Motivated by the strong links between COVID-19 contagion and digital finance, we have proposed a Poisson autoregressive model that can be useful not only to estimate the contagion curve, and monitor its evolution, for public health purposes, but also to predict the financial market dynamics, in the countries affected by contagion.

Future research may involve the construction of a causal model to predict financial market indicators, based on contagion counts, also taking into account the interdependence between different country markets.

Furthermore, a model with time varying parameters, or a non-parametric model (see, e.g., Härdle et al. 2004; West and Harrison 1997), could better capture the contagion dynamics.
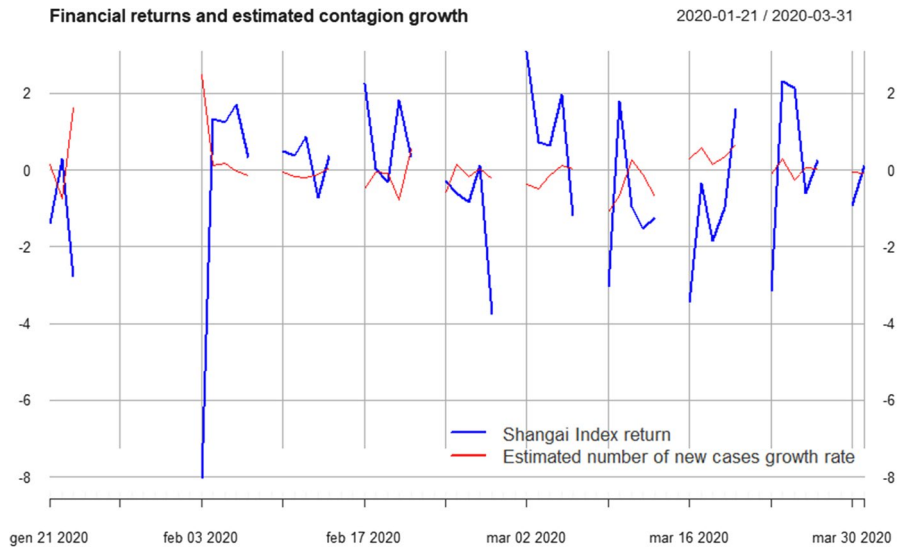
Fig. 5 Comparison between SSE returns and fitted contagion counts variations (the SSE returns were multiplied by 100 to ensure proper comparison of the series). The series are interrupted by market closure days

We leave to further research these methodological extensions, which would also need a longer time series that, at the moment, is not available.

References

Agosto, A., Cavaliere, G., Kristensen, D., & Rahbek, A. (2016). Modeling corporate defaults: Poisson autoregressions with exogenous covariates (PARX). *Journal of Empirical Finance*, *38*(B), 640–663.

Biggerstaff, M., Cauchemez, S., Reed, C., Gambhir, M., & Finelli, L. (2014). Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: A systematic review of the literature. *BMC Infectious Diseases*, *4*(14), 480.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica*, *50*, 987–1008.

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), 1–50.

Ferland, R., Latour, A., & Oraichi, D. (2006). Integer-valued GARCH processes. *Journal of Time Series Analysis*, *27*, 923–942.

Fokianos, K., & Tjøstheim, D. (2011). Log-linear Poisson autoregression. *Journal of Multivariate Analysis*, *102*, 563–578.

Härdle, W., Müller, M., Sperlich, S., & Werwatz, A. (2004). *Nonparametric and semiparametric models*. Berlin: Springer.

Mihoci, A., Althof, M., Chen, C. Y. H., & Härdle, W. K. (2020). FRM financial risk meter. In *Advances in Econometrics, The Econometrics of Networks* (Vol. 42).

Pediroda, V. (2020). Too many companies at liquidity risk: The default probability doubles, 24 Ore, 18 March 2020.

West, M., & Harrison, J. (1997). *Bayesian forecasting and dynamic models*. Berlin: Springer.

World Health Organisation. (2020). Novel Coronavirus (2019-nCoV) situation reports (pp. 1–49).

Yu, L., Härdle, K.W., Borke, I., & Benschop, T. (2019). An AI approach to measuring financial risk. *The Singapore Economic Review*.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.