Mathematical and Statistical Methods for Actuarial Sciences and Finance

Marco Corazza · Cira Perna ·
Claudio Pizzi · Marilena Sibillo
Editors

Mathematical and Statistical Methods for Actuarial Sciences and Finance

MAF 2022

Springer

Editors
Marco Corazza
Department of Economics
Ca' Foscari University of Venice
Venice, Italy

Claudio Pizzi
Department of Economics
Ca' Foscari University of Venice
Venice, Italy

Cira Perna
Department of Economics and Statistics
University of Salerno
Fisciano, Salerno, Italy

Marilena Sibillo
Department of Economics and Statistics
University of Salerno
Fisciano, Italy

ISBN 978-3-030-99637-6 ISBN 978-3-030-99638-3 (eBook)
https://doi.org/10.1007/978-3-030-99638-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume is a collection of papers selected and peer reviewed from the more than 100 presented at the International Conference on Mathematical and Statistical Methods for Actuarial Sciences and Finance–MAF2022, held at the University of Salerno from 20 to 22 April 2022.

In its organizational phase, the course of the COVID-19 pandemic was still unpredictable, and the MAF2022 steering committee made the decision to hold the event in a hybrid form, online or in-person, leaving each participant free to choose the most appropriate mode of participation. Nevertheless, we have always hoped to have the widest possible participation in presence, both as a desired sign of normality and as a return to the tradition of the conference, always characterized by cultural and human exchanges that only in presence can be fully realized.

This year's conference, organized by the Department of Economics and Statistics of the University of Salerno with the collaboration of the Department of Economics of the University of Venice Cà Foscari, is the tenth in a two-year series that began in 2004.

It was in fact in 2003 that the mathematicians and statisticians of the Department of Economics and Statistics of the University of Salerno, colleagues and friends among them, conceived and grew the purpose of developing through scientific meetings the cooperation and exchange of ideas among those who, like them, were engaged in research in actuarial science and finance. The enthusiasm about the initiative was always based on the deep conviction that this interaction would surely bear good fruit.

And so, the initiative has followed regularly, availing since 2008 of the valuable collaboration of the Department of Economics of the University of Venice Cà Foscari.

The first six editions were held in Italy, namely in 2004 and 2006 in Salerno, in 2008 in Venice, in 2010 in Ravello (Salerno), in 2012 again in Venice and in 2014 in Vietri sul Mare (Salerno). The international dimension of the conference has grown over time, attracting a wider and wider audience. Thus, in 2016 the MAF was held in Paris and in 2018 in Madrid. The 2020 edition, already suffering from the COVID-19 pandemic, was held in a fully online version from Venice.

This tenth edition confirms the growing interest of the international scientific community towards the initiative, with about 200 participants, more than 170 scientific contributions proposed in the form of abstracts or papers and four prestigious plenary speakers, namely

Prof. Elsa Fornero, Honorary Professor, University of Turin, who presents an invited talk entitled: "Reform, Inform, Educate": a new paradigm for the sustainability of pension system;
Prof. Massimiliano Caporin, University of Padua, who presents an invited talk entitled: Realized Covariance Modelling, Forecast Error Variance Decompositions and a Model-Based Diebold-Yilmaz Index;
Prof. Marcello Galeotti, University of Florence, who presents an invited talk entitled: Applications of Game Theory to Risk Models: Evolutionary and Cooperative Approaches;
Dr. Michel Dacorogna, Prime Re Solutions, Zug, Switzerland, who presents an invited talk entitled: Pro-Cyclicality Beyond Business Cycles: The Case of Traditional Risk Measurements.

Since 2006, all editions of the conference have been accompanied by a book published by Springer, a product that has often been counted among the most downloaded on the platform. Also, this tenth edition proposes the associated book, with the aim of offering the selected scientific contributions in a concise form of maximum 6 pages, in which the authors present their idea and the methodology behind its development, providing, when possible, an illustrative application.

The goal is to create a forum for comparison of ideas, topics and research perspectives, which embodies and represents at best the soul of MAF as a place of meeting and scientific exchange.

Several are the research areas to which the papers are dedicated with a focus on applicability and/or applications of the results:

Actuarial models, analysis of high-frequency financial data, behavioural finance, carbon and green finance, credit risk methods and models, dynamic optimization in finance, financial econometrics, forecasting of dynamical actuarial and financial phenomena, fund performance evaluation, insurance portfolio risk analysis, interest rate models, longevity risk, machine learning and soft computing in finance, management in insurance business, models and methods for financial time series analysis, models for financial derivatives, multivariate techniques for financial markets analysis, neural networks in insurance, optimization in insurance, pricing, probability in actuarial sciences, insurance and finance, real-world finance, risk management, solvency analysis, sovereign risk, static and dynamic portfolio selection and management, trading systems.

In its almost twenty years, the initiative has always availed itself of the support of the Departments of Economics and Statistics of the University of Salerno (Italy) and of the Department of Economics of the University Ca' Foscari of Venice (Italy) and nonetheless of the scientific associations:

- AMASES—Association for Mathematics Applied to Social and Economic Sciences
- SIS—Italian Statistical Society.

Further, we would also like to express our deep gratitude to the members of the scientific and organizing committees and to all the people whose collaboration contributed to the success of the MAF2022 conference. In particular, our heartfelt thanks go to Giovanna Bimonte and Antonio Naimoli, who have worked unstintingly with great enthusiasm and efficiency, continually showing with their work the sharing of the aims of the initiative. We would also like to thank all the participants for their precious and indispensable contribution.

Finally, we are pleased to inform you that the organizational machine is already at work, looking forward to the MAF2024 edition.

April 2022

Marco Corazza
Cira Perna
Claudio Pizzi
Marilena Sibillo

Contents

Metadata of the chapter that will be visualized in SpringerLink

Book Title	Mathematical and Statistical Methods for Actuarial Sciences and Finance
Series Title	
Chapter Title	Estimating Recovery Curve for NPLs
Copyright Year	2022
Copyright HolderName	The Author(s), under exclusive license to Springer Nature Switzerland AG

Author	Family Name	**Rocci**
	Particle	
	Given Name	**Roberto**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Sapienza University
	Address	Rome, Italy
	Email	
Corresponding Author	Family Name	**Carleo**
	Particle	
	Given Name	**Alessandra**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Roma Tre University
	Address	Rome, Italy
	Email	alessandra.carleo@uniroma3.it
Author	Family Name	**Staffa**
	Particle	
	Given Name	**Maria Sole**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	European University
	Address	Rome, Italy
	Email	

Abstract	The objective of the present paper is to propose a new method to measure the recovery performance of a portfolio of non-performing loans (NPLs) in terms of recovery rate and time to liquidate. The fundamental idea is to draw a curve representing the recovery rates during time, here assumed discretized, for example, in years. In this way, the user can get simultaneously information about recovery rate and time to liquidate of the portfolio. In particular, it is discussed how to estimate such a curve in presence of right censored data, i.e. when the NPLs composing the portfolio have been observed in different periods. Uncertainty

about the estimates is depicted trough confidence bands obtained by using the non-parametric Bootstrap. The effectiveness of the proposal is shown by applying the method to a real financial data set about some portfolios of Italian unsecured NPLs taken in charge by a specialized operator.

Keywords (separated by '-')	Recovery rate - Time to liquidate - NPLs - Censored data

Estimating Recovery Curve for NPLs

Roberto Rocci[1], Alessandra Carleo[2](✉), and Maria Sole Staffa[3]

[1] Sapienza University, Rome, Italy
[2] Roma Tre University, Rome, Italy
alessandra.carleo@uniroma3.it
[3] European University, Rome, Italy

AQ1
AQ2

Abstract. The objective of the present paper is to propose a new method to measure the recovery performance of a portfolio of non-performing loans (NPLs) in terms of recovery rate and time to liquidate. The fundamental idea is to draw a curve representing the recovery rates during time, here assumed discretized, for example, in years. In this way, the user can get simultaneously information about recovery rate and time to liquidate of the portfolio. In particular, it is discussed how to estimate such a curve in presence of right censored data, i.e. when the NPLs composing the portfolio have been observed in different periods. Uncertainty about the estimates is depicted trough confidence bands obtained by using the non-parametric Bootstrap. The effectiveness of the proposal is shown by applying the method to a real financial data set about some portfolios of Italian unsecured NPLs taken in charge by a specialized operator.

Keywords: Recovery rate · Time to liquidate · NPLs · Censored data

1 Introduction

Non-Performing Loans (NPLs) are exposures in state of insolvency, that means loans whose collection by banks is uncertain. As Resti and Sironi [9] point out, an effective recovery depends on the characteristics of the exposure, of the counterparty, on macroeconomic and on internal (to the bank) factors. There is a NPL market that offers banks the opportunity to get rid of non-performing loans by selling them to specialized operators who deal with recovery. The main method for determining the value of Non-Performing Loans is that of discounted financial flows, according to which the value of the loans is equal to the sum of the expected income flows, discounted at a rate consistent with the expected unlevered return of the investor and net of the related recovery costs.

In the case of a performing loan, the borrower is expected to pay principal and interest at the agreed deadlines with a high level of probability (one minus the probability of default, generally low). In this case, the uncertainty in the valuation is limited to the determination of the discount rate to consider the general market conditions and the specific risk of the debtor. In the case of Non-Performing Loans, the uncertainty concerns not only the discount rate but also the amount that will be returned and the time of return. In fact, the probability of default is now equal to one, or is in any case

M. Corazza et al. (Eds.): MAF 2022, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pp. 1–7, 2022.
https://doi.org/10.1007/978-3-030-99638-3_64

very high, if the credit is in other categories of impaired loans (unlikely to pay). The valuation methodologies currently used on the market are therefore based primarily on forecast models of the amount of net repayments expected from receivables and related collection times.

The estimation methodology for recovery rate, which we are interested in for *NPLs*, is faced in the more general context of Basel II. It is well known that under internal ratings-based (IRB) approach to determine capital requirements for credit risk, banks are required to estimate the following risk components: probability of default (PD), loss given default (LGD), exposure at default (EAD) and maturity (M). While the estimation of PD is well established, LGD is not so well investigated and still subject to research. Given the borrower has already defaulted, LGD is defined as the proportion of money financial institutions fail to gather during the collection period and, conversely, Recovery Rate (RR) is defined as the proportion of money financial institutions successfully collected. That means $LGD = 1\text{-}\ RR$.

Recovery rate (or LGD) can be estimated using both parametric and non-parametric methods. Mainly, recovery rate is estimated using parametric methods and considering a one-year time horizon. Methods used in literature, among others, are: classical linear regression, regularized regression like Lasso, Ridge, Elastic-net, etc. [7], Beta regression, inflated Beta regression, two-stage model combining a Beta mixture model with a logistic regression [10].

In the case of NPLs, in our opinion, in investigating the recovery process of defaulted exposures the focus must be not only on the recovered amounts, but also on the duration of the recovery process, the so-called time to liquidate (TTL).

Cheng and Cirillo [4] propose a model that can learn, using a Bayesian update in a machine learning context, how to predict the possible recovery curve of a counterpart. They introduce a special type of combinatory stochastic process, based on a complex system of assumptions, referring to a discretization of recovery rates in m levels.

Our purpose is to introduce a particular non parametric method to measure the performance of a NPLs portfolio in terms of recovery rate (RR) and time to liquidate (TTL) jointly, without assuming any particular model and/or discretization of the RR. The idea is to represent the recovery process as a curve showing how the RR is distributed during the time without assuming a particular parametric model. We will also propose a method to estimate such a curve when some data are censored. The plan of the paper is the following. In Sect. 2 we show how the recovery curve is defined, while in Sect. 3 the method of estimation in case of censored data is introduced. In Sect. 4, the effectiveness of the proposal is shown through an application on real data, while some conclusions and final remarks are discussed in Sect. 5.

2 Recovery Rate and Time to Liquidate of a Portfolio

The definition of recovery rate (RR) and time to liquidate (TTL) of a NPLs portfolio is not trivial because the two quantities are strictly connected. Since it is crucial to decide when to measure the RR and TTL – that is when each NPL in the portfolio has been entirely liquidated or after a given period to be defined - the measurement of the RR cannot disregard the measurement of the TTL and vice versa.

First, we note that to measure the TTL when the last NPL has been liquidated could lead to measures highly affected, and biased, by anomalous NPLs with long TTLs and small EAD. It follows that the TTL should be measured when the RR becomes significant. It remains to understand what is "significant". Second, in many cases the user needs a more complete information rather than only two numbers: RR and TTL. It would be better to know how the RR increases during the time. This would also help in choosing at what RR point to measure the TTL according to whatever optimality criterion the operator decides to adopt. For the aforementioned reasons, we decide to measure the behavior of the RR during the time through what we called the "recovery curve". Such a curve is built in the following way.

Let us consider a portfolio of K NPLs. For each of the K NPLs the debt exposure at default is EAD_{k} (exposure at default of the k-th NPL) and the total portfolio exposure $EAD = \sum_{k=1}^{K} EAD_k$. Assume I time intervals (of the delay of payment) from the default (time t_0) to the valuation date (time t_I). Let $p_{k,i}$ be the recovery of the k-th NPL, in the i-th interval (of delay), i.e. $(t_{i-1}, t_i]$, with $k \in \{1, 2, \ldots, K\}$ and $i \in \{1, 2, \ldots, I\}$. The portfolio recovery in time interval i equals $p_i = \sum_{k=1}^{K} p_{k,i}$, that is the total recovery, for all the K debt positions, in the i-th time interval of delay. Consequently, after i time intervals of delay, i.e. by the end of the interval $(t_0, t_i]$ we define $P_i = \sum_{i'=1}^{i} p_{i'}$ as the total portfolio "recovery value until time t_i", i.e. the total recovery, for all the K debt positions, in the first i periods from the default date. We could also define the total recovery$P*_i = \sum_{i'=1}^{1} \mathrm{V}(p_{i'})$, being $\mathrm{V}(p_i)$ the value of p_i capitalized at an appropriate interest rate. In this initial study, we (like many other, i.e. [10]) do not consider the interest because we consider time and recovery rate together and also because the recovery curve, even if lower, would have the same trend. We define also $R_i = P_i/EAD$ as the portfolio "recovery rate until time i", while $r_i = p_i/EAD$ equals the portfolio recovery rate in the i-th time interval. Since $R_i = \sum_{i'=1}^{i} r_{i'}$, $r_i = R_i - R_{i-1}$ and $r_1 = R_1$ we can refer in an equivalent way to R_i or to r_i for $i \in \{2, \ldots, I\}$.

In Table 1 there is an example of portfolio with $K = 4$ debt positions.

Table 1. Portfolio with $K = 4$ debt positions.

k	EAD_k	$p_{k,1}$	$p_{k,2}$	$p_{k,3}$	$p_{k,4}$
1	100	10	0	0	0
2	200	20	15	0	0
3	300	20	25	10	15
4	400	30	35	10	#N/D

We are interested in measuring the portfolio performance in 3 years after default, i.e. $I = 3$ periods of delay. It can be measured in terms of recovery rates until year i as

Table 2. Portfolio ($EAD = 1000$) performance in 3 years ($I = 3$).

i	1	2	3
p_i	80	75	20
P_i	80	155	175
r_i	8.00%	7.50%	2.00%
R_i	8.00%	15.50%	17.50%

We see that, for example, in the first 2 years the portfolio recovers the 15.5% of the total initial exposure: 8% in the first year and 7.5% in the second.

Sometimes the available data are incomplete, in particular, censored, i.e. the $p_{k,i}$ are not available from a certain date on for some k. In our example, this happens in the fourth period for the NPL ($k = 4$). In this case, it is not possible to compute the recovery curve for the fourth interval without further hypotheses. In the next section, we will discuss some of them and how to estimate the recovery curve from the incomplete data.

3 Estimating the Recovery Rate Curve from Censored Data

The estimation of the recovery curve in the presence of censored data is carried out in a way similar to the estimation of a survival curve (e.g. [8]). First, we note that sometimes it is interesting to consider the "conditional recovery rate" c_i in each delay period i. Let E_i be the effective portfolio exposure at the beginning of period i

$$E_i = \begin{cases} EAD & i = 1 \\ \sum_{k=1}^{K} \left(EAD_k - \sum_{i'=1}^{i-1} p_{k,I'} \right) & i > 1 \end{cases} \tag{1}$$

that means $E_i = EAD - P_{i\text{-}1}$ with $P_0 = 0$ by convention. The conditional recovery rate is defined as $c_i = p_i/E_i$. In words, it is the recovery rate with respect to the effective portfolio exposure at the beginning of the period (E_i) rather than to the initial one (EAD). We observe that it is possible to obtain r_i from c_i and R_{i-1}:

$$\mathrm{r_i} = \frac{p_i}{EAD} \cdot \frac{E_i}{E_i} = \frac{p_i}{E_i} \cdot \frac{EAD - P_{i-1}}{EAD} = c_\mathrm{i}\left(1 - \frac{P_{i-1}}{EAD}\right) = c_i(1 - R_{i-1}) \tag{2}$$

It means that the recovery rate is the conditional recovery of the percentage of how much still has to be recovered. This way of computing r_i is convenient when there are censored data in the database, i.e. for some NPLs the recovery $p_{k,i}$s are observed only until a particular time. In this case, since $r_i = p_i/EAD$ cannot be used, the idea is to apply formula (2) by computing the conditional recovery rate c_i using only the available data. In details, let us suppose that $K_i = \{k = 1, \ldots, K \mid \exists\, p_{k,i}\}$ is the subset of indexes k corresponding to the NPLs for which at delay time i the value $p_{k,i}$ is not censored. In this case the effective portfolio exposure, for $i > 1$, is a generalization of (1):

$$E_i = \sum_{k \in K_\mathrm{i}} \left(EAD_k - \sum_{i'=1}^{i-1} p_{k,i'} \right) \tag{3}$$

and the conditional recovery rate is

$$c_i = \left(\sum\nolimits_{k \in K_i} p_{k,i}\right) / E_i \tag{4}$$

Let's consider the previous example. If we want to consider more than 3 intervals of delay, assuming we are interested in measuring the performances in 4 years, i.e. $I = 4$ periods of delay, then we obtain the same results of Table 2 for the first 3 years, and for $i = 4$ we get: $p_4 = 15$, $P_4 = 190$, $E_4 = 500$, $r_4 = 2.48\%$, $c_4 = 3.00\%$, $R_4 = 19.98\%$. This method of measuring performances allows not only to measure jointly the recovery rate and the time to liquidate, but also to deal with censored data.

Obviously, it is wrong to imagine the censored data equal to 0, meaning no inflows instead than no information about that inflow. With the same example, substituting $p_{4,4} = 0$, we would obtain the same results of Table 2 for the first 3 years, but for $i = 4$ we would get: $p_4 = 15$, $P_4 = 190$, $E_4 = 825$, $r_4 = 1.50\%$, $c_4 = 1.82\%$, $R_4 = 19.00\%$. That is, probably, an underestimate of the true curve.

The results would have been different if we simply did not consider in the portfolio the NPLs for which the data are censored. In the example, considering $I = 4$ periods of delay excluding NPL_4 would lead to different results for all the durations, as it is shown in the table below. Such estimates are of lower quality than the proposed ones because obtained by using less data, i.e. information (Table 3).

Table 3. Portfolio ($EAD = 600$) performance for $K = 3$ loans.

i	1	2	3	4
p_i	50	40	10	15
r_i	8.33%	6.67%	1.67%	2.50%
R_i	8.33%	15.00%	16.67%	19.17%

4 Application

We analyse a data set of Italian NPLs supplied by a specialized operator. We examine two portfolios of unsecured loans with different initial debt size. The portfolios have the same year of acceptance by the operator: year 2005. In particular:

Portfolio 1: $5000 < EAD_k < 25000$, $K = 4732$, Average $EAD_k = 14709$;
Portfolio 2: $100000 < EAD_k < 250000$, $K = 876$, Average $EAD_k = 151117$.

We consider as time t_0 the year of acceptance (2005), rather than the exact time of default, because this is the moment in which the operator starts the recovery procedure. We follow the recovery history for 9 years. We observe that both portfolios have approximately 5% censored data in the last year and about 2.5‰ censored data in the penultimate year. The results in terms of r_i and R_i are reported in the plots below,

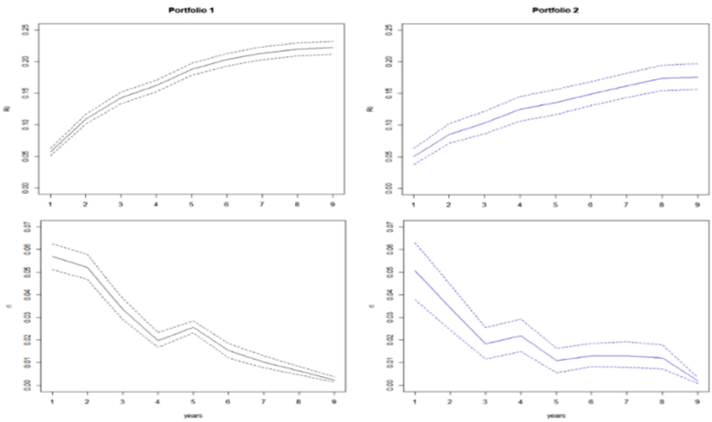

Fig. 1. Recovery rate until time i (R_i) recovery rate (r_i) of Portfolio 1 and Portfolio 2.

where the dotted lines are the boundaries of the confidence intervals computed pointwise by using a non-parametric bootstrap [6].

Obviously, the highest values of the recovery rate are at the beginning of the period ($i = 1$) and as time passes the recovery rate tends to decrease, even if not monotonically. To compare the results we discuss R_i, that in our opinion is the most explicative ratio. Even considering the width of the confidence intervals, it appears that the recovery is greater for the portfolio with smaller credits. Probably, this is because taking charge by specialized operators has greater effect on those who must return lower amounts.

In the extended version of the paper other comparisons will be presented.

5 Conclusions and Final Remarks

According to the objective of this paper, we propose a kind of measurement that takes in consideration both the recovery rate, the time to liquidate and how they interact. This is obtained by estimating a "recovery curve" displaying the behaviour of the recovery rate during the time.

In doing that, we faced the problem of censored data and we suggest to use a method of measuring performances that allows not only to measure jointly the recovery rate and the time to liquidate, but also to deal with censored data. This method is based on an algorithm that is usually used in the construction of survival curves.

Our next goal is to use our method to compare performance of portfolios with different characteristics by using non-parametric boostrap tests for clustered observations. taking into account. Another idea is to extend and test the validity of the method to cases where the database has missing data not only at the end of the observation period, but also at the beginning of it.

References

1. BIS: The Internal Ratings-based Approach. Bank for International Settlements, Basel (2001). https://www.bis.org/publ/bcbsca05.pdf

2. Bruche, M., Gonzales-Aguado, C.: Recovery rates, default probabilities, and the credit cycle. J. Bank. Finance **34**(4), 754–764 (2010)
3. Campbell, J.Y., Lo, A.W., MacKinlay, C.A.: The Econometrics of Financial Markets. Princeton University Press, Princeton (1997)
4. Cheng, D., Cirillo, P.: A reinforced urn process modeling of recovery rates and recovery times. J. Bank. Finance **96**, 1–17 (2018)
5. Devjak, S.: Modeling of cash flows from nonperforming loans in a commercial bank. Naše Gospodarstvo/Our Economy **64**(4), 3–9 (2018)
6. Efron, B.: Censored Data and the Bootstrap. J. Am. Stat. Assoc. **76**(374), 312–319 (1981)
7. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer, New York (2009). https://doi.org/10.1007/978-0-387-84858-7
8. Kalbfleisch, J.D., Prentice, R.L.: The Statistical Analysis Failure Time Data. Wiley, New York (2002)
9. Resti, A., Sironi, A.: Risk Management and Shareholders' Value in Banking. Wiley, New York (2007)
10. Ye, H., Bellotti, A.: Modelling recovery rates for non-performing loans. Risks **7**(1), 19 (2019)

Author Queries

Chapter 64

Query Refs.	Details Required	Author's response
AQ1	This is to inform you that corresponding author has been identified as per the information available in the Copyright form.	
AQ2	Please confirm if the corresponding author email id is correctly identified. Amend if necessary.	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	• • • under matter to remain	
Insert in text the matter indicated in the margin		New matter followed by or
Delete	through single character, rule or underline or through all characters to be deleted	or
Substitute character or substitute part of one or more word(s)	through letter or through characters	new character or new characters
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert ‘superior’ character	through character or where required	or under character e.g. or
Insert ‘inferior’ character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	through character or where required	
Reduce space between characters or words	between characters or words affected	