Impact of COVID-19 on the Robustness of the Probability of Default Estimation Model

Ming-Chin Hung [1],*[ID]**, Yung-Kang Ching** [2][ID] **and Shih-Kuei Lin** [3]

1   Department of Financial Engineering and Actuarial Mathematics, Soochow University, Taipei 100, Taiwan
2   Risk Management Development, China Development Financial Holding, Taipei 105, Taiwan;
    stevenching0729@gmail.com
3   Department of Money and Banking, National Chengchi University, Taipei 116, Taiwan; square@nccu.edu.tw
*   Correspondence: nhungg@scu.edu.tw

Abstract: Probability of default (PD) estimation is essential to the calculation of expected credit loss under the Basel III framework and the International Financial Reporting Standard 9. Gross domestic product (GDP) growth has been adopted as a key determinant in PD estimation models. However, PD models with a GDP covariate may not perform well under aberrant (i.e., outlier) conditions such as the COVID-19 pandemic. This study explored the robustness of a PD model with a GDP determinant (the test model) in comparison with that of a PD model with a credit default swap index (CDX) determinant (the alternative model). The test model had a significantly greater ratio of increase in Akaike information criterion than the alternative model in comparisons of the fit performance of models including 2020 data with that of models excluding 2020 data (i.e., that do not cover the COVID-19 pandemic). Furthermore, the Cook's distance of the 2020 data of the test model was significantly greater than that of the alternative model. Therefore, the test model exhibited a serious robustness issue in outlier scenarios, such as the COVID-19 pandemic, whereas the alternative model was more robust. This finding opens the prospect for the CDX to potentially serve as an alternative to GDP in PD estimation models.

Keywords: Anscombe's quartet; Cook's distance; default rate; expected credit loss; gross domestic product

1. Introduction

Credit risk is generally understood as the potential that a borrower or counterparty will fail to meet its contractual obligations. Banks need to estimate the probability of such events occurring and set aside capital to absorb contingent losses. Loan loss provision estimates are constantly updated based on the bank's potential customer defaults. These estimates are usually calculated based on a probability of default (PD) model, as applied to historical default data. Credit risk evaluation is crucial not only for internal credit decisions but also for regulatory purposes (BCBS [1,2]). In July 2014, the International Accounting Standards Board (IASB) issued the final version of the International Financial Reporting Standard 9 (IFRS 9)—Financial Instruments. The IFRS 9 introduced an expected credit loss (ECL) framework concerning how banks should recognize and manage potential credit losses for financial statement–reporting purposes. IFRS 9 defines principles but grants freedom in choosing which models and approaches banks use to estimate their potential losses. These estimates are then used to determine how much capital is to be set aside as buffers against loss. This ECL practice is aligned with internal ratings–based regulatory practices for determining financial institutions' regulatory capital requirements in Basel III.

Effective from 1 January 2018, the IFRS 9 mandates for the measurement of impairment loss allowances to be based on a forward-looking ECL accounting model rather than on an incurred loss accounting model. The ECL model, which incorporates current and predicted macroeconomic factors, such as expectations in changes in the GDP growth rate,

is designed to yield more accurate predictions of credit losses. As a standard of financial reporting purposes, however, the ECL model can result in volatile credit loss estimation following unexpected events, such as the COVID-19 pandemic.

The income information statement from the Barclays Group 2nd Quarter Financial Report provides an example of this (Figure 1). It stated that a total of £1097 million in profit in Q419 decreased to £359 million in Q220. Conversely, the credit impairment charge increased more than fourfold, from £523 million, in Q419, to £2115 million, in Q120. In Q220, it remained substantially higher than that at Q419, at £1623 million. After Q220, the amount accrued decreased back to £608 million in Q320 and £492 million in Q420. Moreover, due to the economic recovery from the COVID-19 pandemic, £797 million of credit impairment was released in Q221 and the single-quarter profit thus increased to £2580 million. The point-in-time characteristics of forward-looking ECL estimation resulted in great volatility in the bank's profit data.

Barclays Group

Income statement information	Q221 £m	Q121 £m	Q420 £m	Q320 £m	Q220 £m	Q120 £m	Q419 £m	Q319 £m
Net interest income	2,052	1,851	1,845	2,055	1,892	2,331	2,344	2,445
Net fee, commission and other income	3,363	4,049	3,096	3,149	3,446	3,952	2,957	3,096
Total income	**5,415**	**5,900**	**4,941**	**5,204**	**5,338**	**6,283**	**5,301**	**5,541**
Credit impairment releases/(charges)	797	(55)	(492)	(608)	(1,623)	(2,115)	(523)	(461)
Net operating income	**6,212**	**5,845**	**4,449**	**4,596**	**3,715**	**4,168**	**4,778**	**5,080**
Operating costs	(3,587)	(3,545)	(3,480)	(3,391)	(3,310)	(3,253)	(3,308)	(3,293)
UK bank levy	—	—	(299)	—	—	—	(226)	—
Litigation and conduct	(66)	(33)	(47)	(76)	(20)	(10)	(167)	(1,568)
Total operating expenses	**(3,653)**	**(3,578)**	**(3,826)**	**(3,467)**	**(3,330)**	**(3,263)**	**(3,701)**	**(4,861)**
Other net income/(expenses)	21	132	23	18	(26)	8	20	27
Profit before tax	**2,580**	**2,399**	**646**	**1,147**	**359**	**913**	**1,097**	**246**
Tax charge	(263)	(496)	(163)	(328)	(42)	(71)	(189)	(269)
Profit/(loss) after tax	**2,317**	**1,903**	**483**	**819**	**317**	**842**	**908**	**(23)**
Non-controlling interests	(15)	(4)	(37)	(4)	(21)	(16)	(42)	(4)
Other equity instrument holders	(194)	(195)	(226)	(204)	(206)	(221)	(185)	(265)
Attributable profit/(loss)	**2,108**	**1,704**	**220**	**611**	**90**	**605**	**681**	**(292)**

Figure 1. Barclays Group 2nd quarter income statement information at 30 June 2021 [3]. Data source: Barclays Group Quarterly result summary of 2nd quarter Financial report at 30 June 2021.

Barclays' credit cost percentage, which is defined as the credit impairment cost divided by total income, is illustrated in Figure 2. Between Q319 and Q221, the credit cost percentage changed considerably, with a high of 34% ($\approx 2115/6283$), in Q120, and a low of -15% ($\approx -797/5415$), in Q221. The ECL method was designed to improve the accuracy of credit cost predictions. However, when one considers the accounting principle of matching cost with revenue, excessive change in the credit cost percentage can confound inter-period analysis and confuse investors.
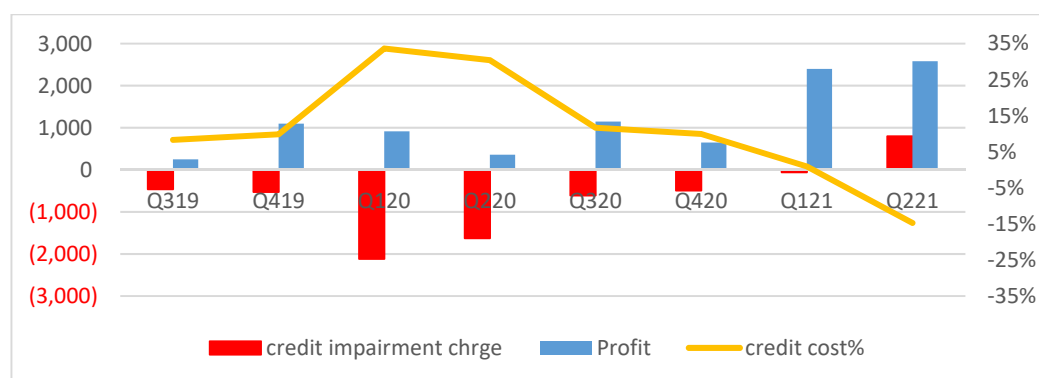
Figure 2. Barclays Group credit cost percentage (credit impairment cost divided by total income), Q319 to Q221.

Motivated by observing the fluctuated credit impairment estimation and credit cost percentage exhibited in Figures 1 and 2, the main aim of this paper is to explore the robustness of the PD model with a GDP determinant. ECL is generally calculated as the PD-weighted average of credit losses, specifically ECL = PD × (exposure at default) × (loss given default). A key factor generally adopted in a PD model is GDP growth. However, this approach may lead to dramatic changes in accounting or financial profit and loss and, thus, result in excessive fluctuations in ECL estimates. As noted, this phenomenon was especially evident during the COVID-19 pandemic in 2020. In this paper, we also explore the usage of a credit default swap index (CDX) determinant in a PD model in place of a GDP determinant to reach a less volatile credit loss estimation. The remainder of this paper is structured as follows: Section 2 reviews the literature on the PD model. Section 3 describes the empirical results of PD models with versus without the COVID-19 data. Section 4 concludes the paper.

2. Literature Review

The Basel III framework and the IFRS 9 were introduced following the financial crisis and European debt crisis. Basel III regulates bank capital, whereas IFRS 9 specifies how banks should classify their assets and estimate their future credit losses. Under IFRS 9, as a part of lifetime ECL calculations for stage 2 credit assets, banks must estimate multiperiod lifetime PDs. Under the Basel accord, PDs are commonly estimated as through-the-cycle to neutralize economic fluctuations and achieve lower volatility in credit risk capital requirements. Conversely, under the IFRS 9, PD estimates should be point-in-time and include forward-looking information, especially for macroeconomic forecasts [4,5]. The COVID-19 pandemic is the first stressful economic scenario since the implementation of IFRS 9 in 2018. In this study, we focus on the effect of COVID-19 on the robustness of PD estimation.

2.1. Logistic Regression PD Model

PDs are of interest to practitioners in financial institutions, as well as to regulators. Logistic regression has been widely adopted for PD models because of its simplicity and amenability to intuition and explanation (Crook et al. [6]). For example, one of the most popular credit risk models is the credit portfolio view model, which is analyzed using logistic regression and contains macroeconomic factors, such as the GDP, as the systematic explanatory variables.

Logistic regression is a common classification method when the response variable is binary, such as whether a default or nondefault occurs. A sound logistic regression model should feature high interpretability, high predictive power, and robustness to data outliers and default sparsity. Given a binary response variable L and a set of covariates $\boldsymbol{x}$, the basic setup of the logistic regression model is as follows: Conditional on $\boldsymbol{x}$, the response variable L is assumed to be Bernoulli distributed; that is, $L|x \sim Bernoulli\ (p)$ for some $p \in [0,1]$. The goal of logistic regression is to fit a predictive model for the binary response variable. Let the random variable $L_{i,t}$ be defined as:

$$L_{i,t} = \begin{cases} 1, & \text{default in } i^{\text{th}} \text{ loan at year } t \\ 0, & \text{no default in } i^{\text{th}} \text{ loan at year } t \end{cases}$$

and the PD for a rating class in the same year is assumed to be constant. The observation of n_t credit exposures can be written as

$$\boldsymbol{L}_t \equiv (L_{1,t}, \cdots, L_{i,t}, \cdots, L_{n_t,t}) \text{ with } L_{i,t}|\boldsymbol{x}_t \sim \text{Bernoulli}(1; p_t), \ i = 1, \cdots, n_t.$$

In the case of binomial data, the random variables $L_{1,t}, \cdots, L_{i,t}, \cdots, L_{n_t,t}$ are assumed to be independent and Y_t, the number of defaults observed, is defined as:

$$Y_t \equiv L_{1,t} + L_{2,t} + \cdots + L_{n_t,t}.$$

As $L_{1,t}, \cdots, L_{i,t}, \cdots, L_{n_t,t}$ are n_t independent and identically distributed trials, it can be inferred that, conditional on x_t, Y_t follows the binomial distribution

$$Y_t | x_t \sim binomial(n_t; p_t), \quad t = 1, \cdots, T,$$

where T is the number of years in the default data set. By utilizing the logit relationship

$$logit(p_t) = \ln \frac{p_t}{1 - p_t} = \boldsymbol{\beta}' x_t,$$

in terms of the logistic density function, the conditional probability of default number at year t is

$$
\begin{aligned}
P(Y_t = y | x_t; \boldsymbol{\beta}) &= \binom{n_t}{y} p_t^y (1 - p_t)^{n_t - y} \\
&= \binom{n_t}{y} \left(\frac{1}{1 + e^{-\boldsymbol{\beta}' x_t}} \right)^y \left(1 - \frac{1}{1 + e^{-\boldsymbol{\beta}' x_t}} \right)^{n_t - y}
\end{aligned}
\tag{1}
$$

The likelihood function, assuming that all the observations $(Y_1, Y_2, \cdots, Y_T)$ are independent and binomially distributed, is defined as

$$Lik(\boldsymbol{\beta} | y; x) = \prod_{t=1}^{T} P(y_t | x_t; \boldsymbol{\beta}),$$

and the log-likelihood function that is defined as

$$\ell(\boldsymbol{\beta} | y; x) \equiv \log Lik(\boldsymbol{\beta} | y; x) = \sum_{t=1}^{T} P(y_t | x_t; \boldsymbol{\beta}) \tag{2}$$

is maximized using various optimization techniques, such as the gradient descent method. Furthermore, the associated Akaike information criterion (AIC) is defined as

$$AIC = 2k - 2 \log \hat{Lik}(\cdot), \tag{3}$$

where k is the number of model parameters and $\log \hat{Lik}(\cdot)$ is the maximum value of the log-likelihood function in Equation (2). As the equation expresses a property of the penalty (negative) function, a smaller AIC value suggests a better fit.

2.2. PD Model with a GDP Determinant

This study primarily aimed to explore the robustness of a PD model with a GDP covariate versus that of a PD model with a CDX covariate. Estimating PDs is challenging due to the limited availability of data and the sparsity of defaults. We explored the logit model with GDP growth as a macroeconomic parameter. The relationship between PD and various macroeconomic variables has been modeled in many applications. Most of the papers discussed in this subsection have demonstrated that GDP growth is significantly related to the default rate (DR).

For the banking sector, Jabra [7] used a binomial logit model and demonstrated how much bank default in the European banking system can be explained, not only by CAMELS (capital adequacy, asset quality, management quality, earnings potential, liquidity, and sensitivity to market risk) variables, but also by GDP growth. Bonjini et al. [8] discovered that bank defaults in developing countries increase with the severity of macroeconomic shocks. Arena [9] and Männasoo and Mayes [10] demonstrated that increased GDP growth (as a macroeconomic indicator) significantly reduced bank PD. Ortolano and Angelini [11] noted that the highest correlation between PD and various adopted GDP covariates was -24%. The negative value corroborated the finding of a relationship between GDP and banking credit risk assessment reported by Jabra, Mighri, and Mansouri [12].

With regard to the corporate sector, Simons and Rolwes [13] provided robust evidence for a relationship between macroeconomic variables and GDP growth in the default behavior of Dutch firms. This observation led to the implementation of econometric models that

describe PD in terms of macroeconomic variables. Couderc and Renault [14] demonstrated that GDP growth is a significant macroDR of firms listed in the Standard and Poor's 500 index and Jakubik [15] demonstrated the same for Finnish firms in 1988–2003 in addition to reporting that interest rate was a nonsignificant macroeconomic variable. By studying the relationship between the credit cycle and macroeconomic variables using data on the rating changes and defaults of US corporations, Koopman et al. [16] demonstrated that many of the variables that were conventionally thought to explain the credit cycle were nonsignificant, with the exception of GDP growth. Virolainen [17] used Finnish data and reported a significant relationship between corporate sector DR and several macroeconomic factors, including GDP. Using data on nonfinancial corporate bond DRs over 150 years, Giesecke et al. [18] studied the relationship between credit default and macroeconomic variables and determined that change in GDP is a strong predictor of DR. Penikas [19] reported that default correlation tends to align with systemic factors, such as the GDP growth rate.

2.3. Robustness of a Model

Anscombe [20] constructed four data sets that yielded the exact same linear regression outputs, namely, the number of observations, mean of the independent variable, mean of the dependent variable, estimated regression coefficient, regression sum of squares, residual sum of squares, estimated standard error of the regression coefficient, and coefficient of determination. However, the four data sets had different characteristics due to the presence of various types of outliers. Intuitively, an outlier is an observation that appears to be "different" from other observations in a data set. An outlier can come in one of three forms: (a) outlier with respect to the dependent variable; (b) outlier with respect to the independent variable (a leverage point); and (c) outlier with respect to both the dependent and independent variables. An outlier can be influential or not influential. An influential observation is an observation whose inclusion in the data set would greatly change the analytical result.

To measure the degree of influence the ith data point has on the analytical result, a natural step is to compute the difference in the fitted results when the ith data point is included and when it is excluded. Cook's distance [21] is based on such an idea for a generalized linear model. An approximation of Cook's distance measure of influence has also been also formulated (Fox [22]). Outliers can distort estimates of binary logit models and linear regression models. In this study, logistic regression diagnostics were performed using the statsmodels package [23] in Python [24]. This measure was computed based on a one-step approximation of the results after one observation was deleted. The diagnostic analytics can also be conducted by means of the local influence method [25].

2.4. CDX as a Determinant in the PD Model

All the studies reviewed in Section 2.2 demonstrated a significant negative relationship between DR and GDP growth. However, using an estimation framework presentation of lifetime PD in accordance with the IFRS 9, Đurović [26] reported that the state of the macroeconomy had a small effect on PD development. He argued that PD development is mainly affected by a rapidly changing marketplace and a constant increase in the number of market participants. Obeid [27] examined data from 40 commercial banks in the Arab world and reported a nonsignificant effect of GDP on bank defaults. Chortareasab et al. [28] performed a meta-analysis of 56 empirical studies on the effect of GDP on nonperforming loans. Their results revealed that GDP performance does not have a predictable effect on credit quality.

Using a regime-switching model, Giesecke et al. [18] reported that change in GDP is a strong predictor of DRs. Surprisingly, however, they also reported that credit spreads do not adjust to current DRs or macroeconomic conditions. Conversely, in studying the effect of credit default swap (CDS) spread determinants on the probability of default, Ortolano and Angelini [11] demonstrated that the price of CDSs is a sound indicator of banks'

creditworthiness. By contrast, Collin-Dufresne et al. [29] demonstrated that credit spreads are driven by factors that are difficult to explain using a standard credit model. Fu et al. [30] revealed that firm performance and macroeconomic conditions play a significant role in explaining CDS spreads.

In studying the fit performance of a PD model, Hu et al. [31] concluded that high-rated companies exhibit a greater need to use market-traded information, such as the CDX, to capture changes in the DR. The similarities and differences between this paper and that of Hu et al. [31] are as follows. A logistic regression model was used as the underlying PD model in both papers. In addition, Moody's DR and the IMF's GDP datasets were used in both papers with different time periods. The paper from Hu et al. [31] was motivated by observing the poorly fitted results of the PD model with GDP determinants, whereas this paper is motivated by observing the extreme mismatch between the behaviour of GDP and DR over 2020. The main criteria of model comparison used by Hu et al. [31] were *p* values and AIC, whereas we primarily use Cook's distance and AIC increasing ratio (see Section 3.2). The empirical results registered by Hu et al. [31] related to goodness-of-fit, especially for the companies from higher rated classes, whereas we, mainly, have determined the impact on PD and, thus, ECL estimation through outlier observation. In other words, the results of the PD model (with GDP determinant) revealed a serious lack of robustness in the 2020 data originating from COVID-19.

To explore the robustness of the PD model with a GDP determinant, we compared the fitted results and influence measures of a PD model using a GDP covariate to that of a PD model using a CDX covariate in the following empirical study.

3. Data and Empirical Results

3.1. Data Descriptions

This study used GDP growth data from the International Monetary Fund (IMF) and corporate DRs by letter rating from a data set of Moody's for 2004–2020. The index of investment grade (Baa and higher) credit default swaps (CDX.NA.IG) starts from 2004, whereas the index of high yield (Ba and lower) credit default swaps (CDX.NA.HY) starts from 2006. To compare the PD fit performance, the same data period were used for both the GDP and CDX explanatory variables. In addition, because defaults are rare among the highest-rated credit entities, the analysis was restricted to the rating classes r = [A, Baa, Ba, B].

The descriptive statistics and kernel density estimates of DRs and GDP growth are illustrated in Table 1 and Figure 3. GDP growth, in Table 1 and the left panel of Figure 3, indicates left-skewed distributions with significant negative skew-test statistics and *p* values of 0.001. Conversely, the DRs of all ratings in the right panel of Figure 3 are all positive, right-skew-distributed with *p* values close to zero. This phenomenon aligns with the reverse relationship between macroeconomic factors (such as GDP) and DRs.
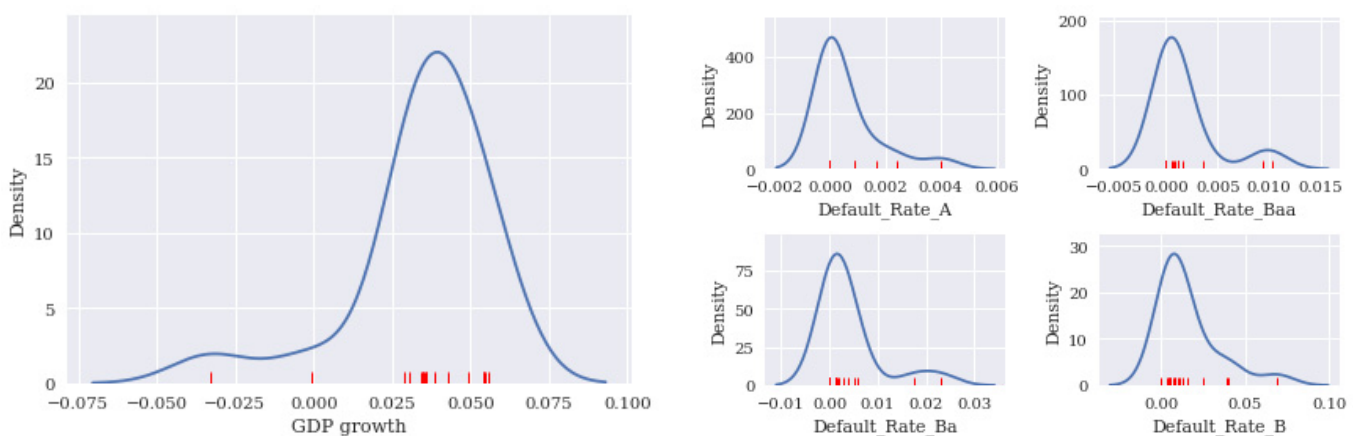

Figure 3. Kernel density estimates of GDP growth rate (**left panel**) and default rates (**right panel**) for ratings A, Baa, Ba, and B.

**Table 1.** Descriptive statistics of GDP and default rates for ratings A, Baa, Ba, and B.

Rating (GDP)	A [1]	Baa	Ba	B	GDP-gr [2]
Count	17	17	17	17	17
Mean	0.0006	0.0018	0.0040	0.0157	0.0344
Std	0.0011	0.0032	0.0065	0.0180	0.0220
Min	0.0000	0.0000	0.0000	0.0000	−0.0330
25%	0.0000	0.0000	0.0000	0.0049	0.0343
(Median)	0.0000	0.0007	0.0014	0.0081	0.0359
75%	0.0009	0.0012	0.0038	0.0157	0.0491
Max	0.0040	0.0103	0.0232	0.0687	0.0556
Skew	2.20	2.26	2.38	2.00	−2.08
Skew_test	3.42	3.49	3.62	3.19	−3.28
p-value of skew test	0.0006	0.0005	0.0003	0.0014	0.0010

Data sources: [1] Moody's Default Reports [32]; [2] IMF, real GDP growth [33].

As illustrated in Figure 4, GDP (red line) was slightly negatively correlated with DR. Two of the most liquid CDX indices were the CDX.NA.IG and CDX.NA.HY. The daily data of the CDX.NA.IG (since 2004) and CDX.NA.HY (since 2006) are displayed in Figure 5, and the trends for the two indexes run in opposite directions. This was because high-yield CDS indices (dotted red line) are conventionally quoted in prices, whereas investment-grade equivalents are quoted in spread basis points.
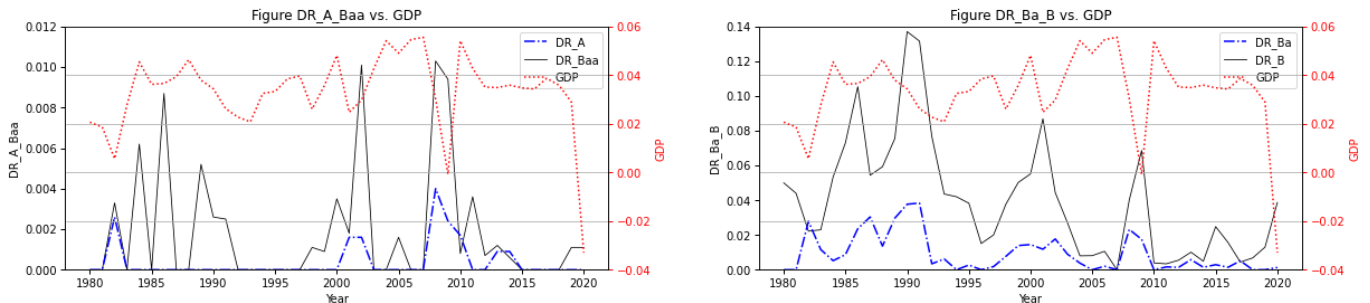
Figure 4. Time series of GDP (see **right** y-axis) and DRs (see **left** y-axis) for ratings A and Baa (**left panel**) and for ratings Ba and B (**right panel**).
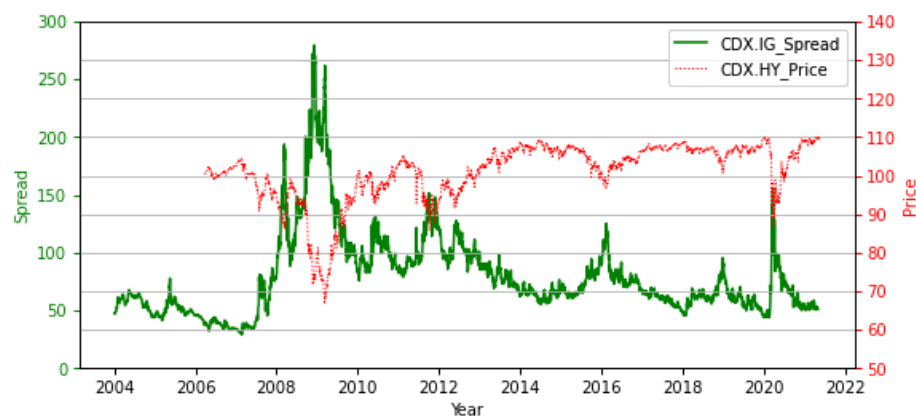
Figure 5. CDX.NA.IG (spread) and CDX.NA.HY (price). Data sources: Datastream database and Bloomberg.

3.2. Empirical Results

Figures 6–9 present the fitted curves (left panels) and Cook's distances (right panels) of the binomial logistic regression for ratings A, Baa, Ba, and B, respectively. For both the left and right panels in Figures 6–9, the left subplots (a) and (c) illustrate the PD model with the GDP covariate (the test model), whereas the right subplots (b) and (d) illustrate the PD model with the CDX covariate (alternative model). To explore the robustness of PD estimation following the COVID-19 pandemic in 2020, we fitted the PD model and calculated the AIC (displayed in the left panel as a legend) and influence measures (displayed in the right panel) for the regression of DR on GDP and on CDX with two data sets, both excluding the 2020 data point (subplots (a) and (b)) and including the 2020 data point (subplots (c) and (d)). We fitted the DRs of the investment rating classes [*A*, *B*aa] on the CDX.NA.IG index and fit the DRs of the noninvestment ratings [*B*a, *B*] on the CDX.NA.HY index. For demonstrative purposes and due to data limitations, we used the CDX annual average to fit Moody's annual DRs.

Starting from the probability density function in Equation (1), we used maximum likelihood estimation (MLE) and the associated AIC in Equation (3) to select the best-fitting logistic regression model for historical DRs. The results were obtained using MLE and the expectation–maximization algorithm, as implemented in the statsmodels fitting procedure in Python.

3.2.1. Comparison of Fitted Curves in the Left Panels and Cook's Distance in the Right Panels

As evident in the left panels of Figures 6–9, the fitted curve in subplot (a) flattened to a near-horizontal curve in subplot (c). In other words, compared with the fit in subplot (a) of Figures 6–9, the marked 2020 data point was dominant and flattened the fitted curve in the lower-left subplot (c). This demonstrates that the PD model with a GDP determinant was not robust for data after the COVID-19 pandemic. The 2020 data point, illustrated in the left panels of Figures 6–9, was an outlier due to its negative value (−3.3%) and was an influential observation (having a high leverage point). This phenomenon occurs in one of the outlier cases in Anscombe's quartet. Conversely, such non-robustness was not evident in the alternative models illustrated in subplots (b) and (d).
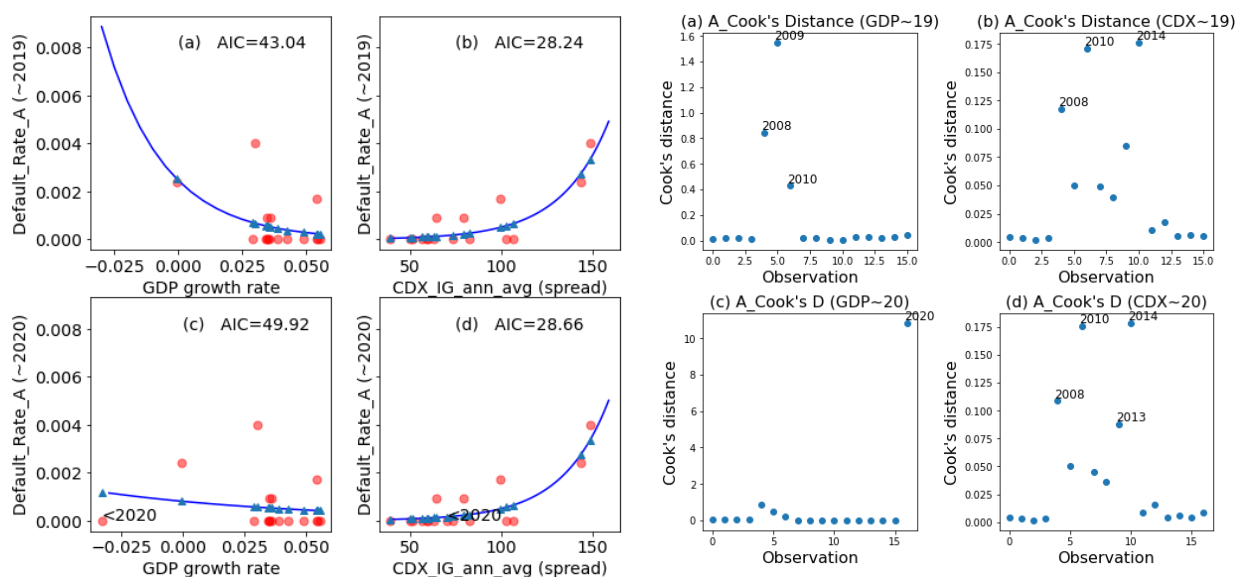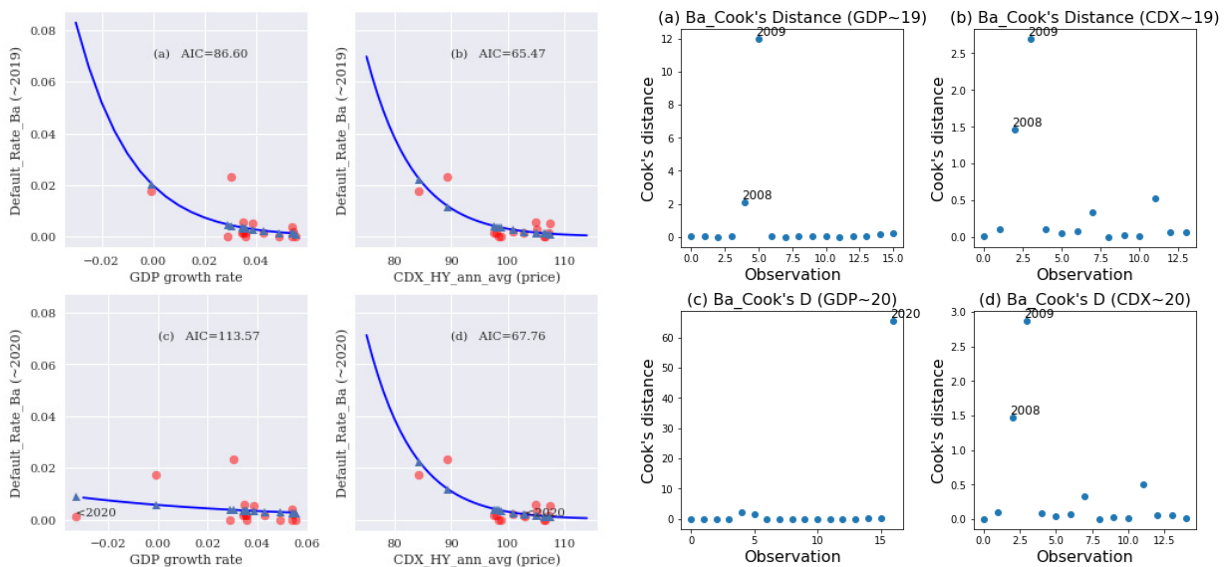

Figure 6. Fitted logistic regression models (**left panel**) and Cook's distances (**right panel**) of default rate in relation to (**a**) GDP in 2004–2019 data; (**b**) CDX.NA.IG in 2004–2019 data; (**c**) GDP in 2004–2020 data; and (**d**) CDX.NA.IG in 2004–2020 data (spread) for A class ratings.
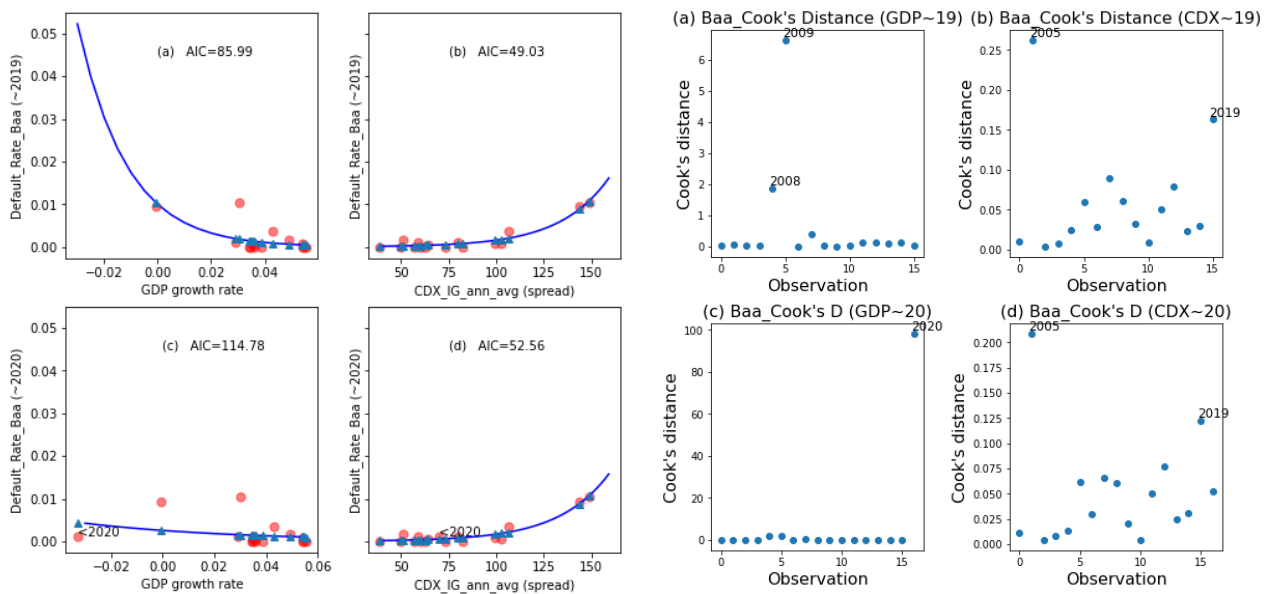
Figure 7. Fitted logistic regression models (**left panel**) and Cook's distances (**right panel**) of default rate in relation to (**a**) GDP in 2004–2019 data; (**b**) CDX.NA.IG in 2004–2019 data; (**c**) GDP in 2004–2020 data; and (**d**) CDX.NA.IG in 2004–2020 data (spread) for Baa class ratings.
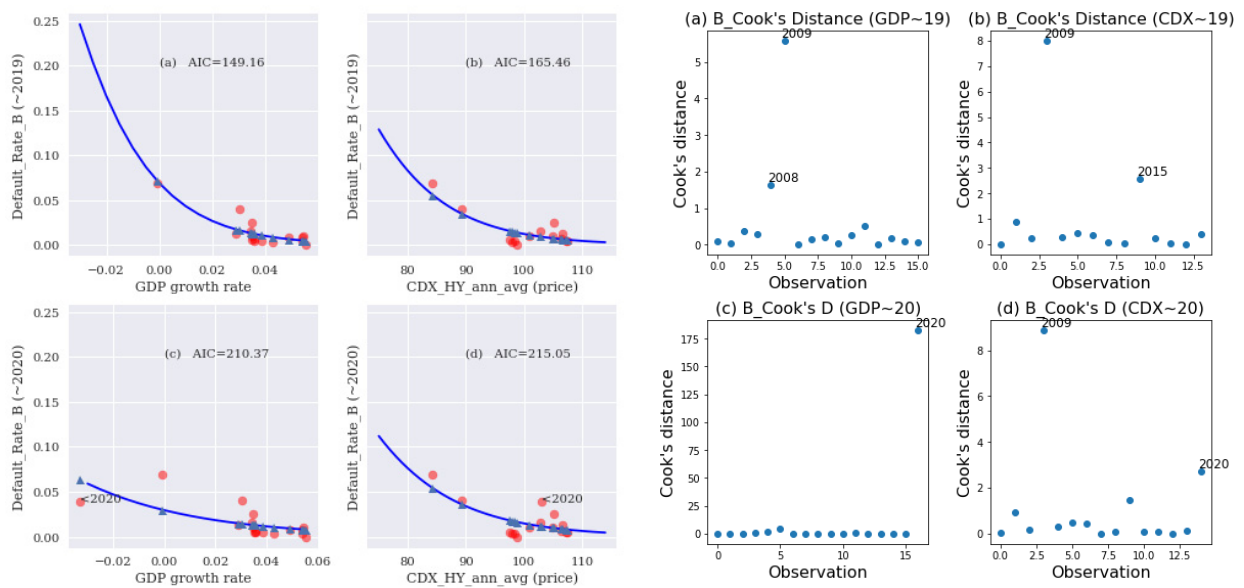

Figure 8. Fitted logistic regression models (**left panel**) and Cook's distances (**right panel**) of default rate in relation to (**a**) GDP in 2006–2019 data; (**b**) CDX.NA.HY in 2006–2019 data; (**c**) GDP in 2006–2020 data; and (**d**) CDX.NA.HY in 2006–2020 data (price) for Ba class ratings.

For Figures 6–9, the fitted curve of subplot (c), in the left panel, echoes the high Cook's distance of that 2020 data point in the upper right corner of subplot (c), in the right panel. Note that, in the right panel of Figures 6–9, the scales substantially differ between subplots (a) and (c). As an example, consider rating A in Figure 6. The y-axis range in subplot (a) for GDP (data through 2019) is narrow at (0, 1.6) in contrast to the wide range of (0, 10) for GDP (data through 2020) in subplot (c). For the test model, the 2020 data point (located in the upper right corner of subplot (c) in the right panel) had an influence value of 11. By contrast, the 2020 data point in subplot (d) for the alternative model only had an influence number of 0.0083, which is not influential in the model fitting process. A similar pattern could be observed for the ratings Baa, Ba, and B.

Figure 9. Fitted logistic regression models (**left panel**) and Cook's distances (**right panel**) of default rate in relation to (**a**) GDP in 2006–2019 data; (**b**) CDX.NA.HY in 2006–2019 data; (**c**) GDP in 2006–2020 data; and (**d**) CDX.NA.HY in 2006–2020 data (price) for B class ratings.

3.2.2. Comparison of Ratio of Increase of AIC

The property of the penalty function implies that a smaller AIC value indicates a better fit. Furthermore, the length of time covered by the data (sample size) affects the calculation of the likelihood function and AIC values. The associated AIC value for the fit of each model is displayed as a legend in each subplot of the left panel in Figures 6–9. To compare the robustness between the test and alternative model following the COVID-19 pandemic, we used the concept of the (negative) rate of return in finance and defined the ΔAIC and $AIC_{increasing}$ ratio as

$$\Delta AIC = AIC_{including\ 2020\ data} - AIC_{excluding\ 2020\ data}$$

and

$$AIC_{ir} = \frac{AIC_{including\ 2020\ data} - AIC_{excluding\ 2020\ data}}{AIC_{excluding\ 2020\ data}}$$

Correspondingly, the more AIC_{ir} increases from the addition of a data point to a data set, the less the robustness of the model with respect to that data point. For example, consider the GDP versus CDX determinant for the A rating in Figure 6. $AIC\_ir\_GDP\_A = (49.92 - 43.04)/43.04 = 0.16$ and $AIC\_ir\_CDX\_A = (28.66 - 28.24)/28.24 = 0.01$, which demonstrate that the alternative model is much more robust than the test model.

Table 2 contains the results for the AIC_{ir} and ΔAIC of the test model versus the alternative model for ratings A, Baa, Ba, and B. Every AIC_{ir} and ΔAIC of the test model, for each rating, was significantly higher than that of the alternative model. That is, the alternative model was more robust than the test model for all ratings, especially for the higher-rated classes, namely A, Baa, and Ba. As the DR of rating B in 2020 remained at the high level of 3.87% (see subplots (c,d) of the left panel of Figure 9), the difference of the ratio of increase in AIC between these two models (0.41 vs. 0.30) was not highly significant relative to the higher-rated classes. This phenomenon may be because the target to buyback from the quantitative easing (QE) policy primarily centers on higher-rated bonds.

Table 2. AIC_{ir} and ΔAIC for the PD model with a GDP determinant versus the PD model with a CDX determinant.

Determinants in PD Model Ratings	AIC_{ir}		ΔAIC	
	GDP	**CDX**	**GDP**	**CDX**
A	0.16	0.01	6.88	0.42
Baa	0.33	0.07	28.79	3.53
Ba	0.31	0.03	26.97	2.29
B	0.41	0.30	61.21	49.59

4. Conclusions and Remarks

Overall, the test model functioned well for normal economic conditions (with data through 2019) but was less robust following the COVID-19 pandemic. The test model had a considerably greater ratio of increase in the AIC than the alternative model in comparisons of fit performance when the 2020 data point (representing the onset of the COVID-19 pandemic) was included versus when it was excluded. Furthermore, the Cook's distance of the 2020 data point of the test model was significantly greater than that of the alternative model. In conclusion, the test model exhibited serious problems with robustness in terms of outliers, such as a global pandemic, especially for high-rated classes, whereas the alternative model was much more robust. These findings echo those of a recent IMF working paper (Roch and Roldán [34]) that examined why countries have issued sovereign state–contingent bonds on only a modest scale and traded them at a large discount, despite the well-known benefits discussed in the literature. They discovered that, for state-contingent bond structures such as the GDP-linked bond issued by Argentina in 2005, a model lacking robustness generates ambiguity premia in bond spreads that are labeled as novelty premia. Their findings rationalize the scarcity of state-contingent debt instruments in practice. A PD model of sovereign default with robustness is required to avoid the novelty premium and increase market liquidity.

The impact of the 2020 data point in this analysis is similar to that of a case introduced in Anscombe's Quartet [20], which indicates that a model fit is predominantly determined by an influential data point. In the present case, we determined that the PD model based on GDP growth was non-robust after the COVID-19 pandemic's commencement on the basis of an additional data point in 2020 (applied to each rating group). However, in the theoretical sense, the 2020 data point involved an observation of a binomial distribution with parameters $(n_t; p_t)$, which was formed from n_t Bernoulli trials (default or nondefault) with n_t as the number of the rated companies in 2020 for each rating class. From a practical perspective, on the other aspect, the ECL estimation occurred on both monthly and quarterly bases (e.g., see Figure 1). However, in this paper, the reported results were based only on the yearly observations due to the data availability constraints of the DRs. The DRs in 2020 were presumably lower than they would have been if governments had not intervened. Therefore, the 2020 DRs may not reflect the true economic situation indicated by the GDP drop. The problem is that whether (and if so, when) government support programs, such as QE, will intervene in the market is unknown.

One limitation of this study is the use of a single-factor model instead of a multi-factor model because our main purpose was to illustrate the robustness of using a GDP versus CDX determinant in a PD model. Furthermore, the analyzed data represented default only as a binary variable (default or nondefault). However, especially within the Basel framework, banks use rating systems with multiple rating grades. Using multiple rating grades would force the adjustment of default probabilities as well as the consideration of the transition probabilities between rating grades. Therefore, future studies can use Markov chains to capture this phenomenon. Furthermore, our PD estimation was based on realized GDP growth. However, the difference between predicted and realized GDP growth may lead to greater fluctuation in the estimation of credit loss. For example,

IMF-predicted 2020 GDP growth was less than −5%, in contrast to the realized −3.3% used in this study. The use of the predicted value in estimating ECL would have caused more serious robustness issue of the PD model. Hence, we argue that some market-based index should be introduced into the PD model. Accordingly, the fluctuating ECL scenario exhibited in Figures 1 and 2 may be at least partially resolved.

References

1. BCBS. *Consultative Document Guidelines—Guidance on Accounting for Expected Credit Losses*; 2015. Available online: https://www.bis.org/bcbs/publ/d350.pdf (accessed on 21 April 2021).
2. Skoglund, J.; Chen, W. The application of credit risk models to macroeconomic scenario analysis and stress testing. *J. Credit. Risk* **2016**, *12*. [CrossRef]
3. Barclays PLC Interim Results Announcement. 2021. Available online: https://home.barclays/content/dam/home-barclays/documents/investor-relations/ResultAnnouncements/H12021/20210728-BPLC-H12021-ResultsAnnouncement.pdf (accessed on 6 July 2021).
4. Novotny-Farkas, Z. The Significance of IFRS 9 for Financial Stability and Supervisory Rules. A Study from the European Parliament's Committee on Economic and Monetary Affairs. 2015. Available online: https://www.europarl.europa.eu/RegData/etudes/STUD/2015/563461/IPOL_STU%282015%29563461_EN.pdf (accessed on 6 July 2021).
5. Ewanchuk, L.; Frei, C. Recent Regulation in Credit Risk Management: A Statistical Framework. *Risks* **2019**, *7*, 40. [CrossRef]
6. Crook, J.N.; Edelman, D.B.; Thomas, L.C. Recent developments in consumer credit risk assessment. *Eur. J. Oper. Res.* **2007**, *183*, 1447–1465. [CrossRef]
7. Jabra, W.B. The fundamental determinants of bank default for european commercial banks. *Int. J. Res. Commer. Manag. Stud.* **2021**, *3*, 1–23.
8. Bongini, P.; Claessens, C.; Ferri, G. The Political Economy of Distress in East Asian Financial Institutions. *J. Financ. Serv. Res.* **2001**, *19*, 5–25. [CrossRef]
9. Arena, M. Bank failures and bank fundamentals: A comparative analysis of Latin America and East Asia during the nineties using bank-level data. *J. Bank. Financ.* **2008**, *32*, 299–310. [CrossRef]
10. Männasoo, K.; Mayes, D.G. Explaining bank distress in Eastern European transition economies. *J. Bank. Financ.* **2009**, *33*, 244–253. [CrossRef]
11. Ortolano, A.; Angelini, E. Do CDS Spread Determinants Affect the Probability of Default? A Study on the EU Banks. Bank I Kredyt 2020. pp. 1–32. Available online: https://bankikredyt.nbp.pl/content/2020/01/BIK_01_2020_01.pdf (accessed on 6 April 2021).
12. Ben Jabra, W.; Mighri, Z.; Mansouri, F. Determinants of European bank risk during financial crisis. *Cogent Econ. Financ.* **2017**, *5*, 2017. [CrossRef]
13. Simons, D.; Rolwes, F. Macroeconomic Default Modeling and Stress Testing. 2009. Available online: https://www.ijcb.org/journal/ijcb09q3a6.pdf (accessed on 6 July 2021).
14. Couderc, F.; Renault, O. Times-to-Default: Life Cycles, Global and Industry Cycle Impact. *FAME Res. Pap. Ser.* **2005**, 142. Available online: https://www.researchgate.net/profile/Olivier-Renault-3/publication/5021538_Times-To-DefaultLife_Cycle_Global_and_Industry_Cycle_Impact/links/5632213208ae13bc6c3779c8/Times-To-DefaultLife-Cycle-Global-and-Industry-Cycle-Impact.pdf (accessed on 6 July 2021).
15. Jakubik, P. *Does Credit Risk Vary with Economic Cycles? The Case of Finland*; IES Working Paper No. 2006/11; Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies: Prague, Czechia, 2006.

16. Koopman, S.J.; Kraussl, R.G.W.; Lucas, A.; Monteiro, A. Credit Cycles and Macro Fundamentals. Tinbergen Institute Discussion Paper. No. 06-023/2. 2006. Available online: https://papers.tinbergen.nl/06023.pdf (accessed on 14 July 2021).
17. Virolainen, K. Macro Stress Testing with a Macroeconomic Credit Risk Model for Finland, Bank of Finland Discussion Papers. 2004. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=622682 (accessed on 20 April 2021).
18. Giesecke, K.; Longstaff, F.A.; Schaefer, S.; Strebulaev, I. Corporate bond default risk: A 150-year perspective. *J. Financ. Econ.* **2011**, *102*, 233–250. [CrossRef]
19. Penikas, H. Why the Conservative Basel III Portfolio Credit Risk Model Underestimates Losses. Available online: http://ceur-ws.org/Vol-2795/paper7.pdf (accessed on 20 April 2021).
20. Anscombe, F.J. Graphs in Statistical Analysis. *Am. Stat.* **1973**, *27*, 17–21. [CrossRef]
21. Cook, R.D. Detection of Influential Observation in Linear Regression. *Technometrics* **1977**, *19*, 15. [CrossRef]
22. Millar, P.; Fox, J. An R and S-Plus Companion to Applied Regression. *Can. J. Sociol. Cah. Can. Sociol.* **2003**, *28*, 110. [CrossRef]
23. Python Statsmodels Package. Available online: https://www.statsmodels.org/stable/index.html (accessed on 25 April 2021).
24. Pregibon, D. Logistic Regression Diagnostics. *Ann. Stat.* **1981**, *9*, 705–724. [CrossRef]
25. Liu, Y.; Mao, G.; Leiva, V.; Liu, S.; Tapia, A. Diagnostic Analytics for an Autoregressive Model under the Skew-Normal Distribution. *Mathematics* **2020**, *8*, 693. [CrossRef]
26. Đurović, A. Macroeconomic Approach to Point in Time Probability of Default Modeling—IFRS 9 Challenges. *J. Cent. Bank. Theory Pr.* **2019**, *8*, 209–223. [CrossRef]
27. Obeid, R. Bank Failure Prediction in the Arab Region Using Logistic Regression Model. 2021. Available online: https://www.researchgate.net/publication/351308263_Bank_Failure_Prediction_in_the_Arab_Region_Using_Logistic_Regression_Model (accessed on 20 April 2021).
28. Chortareasab, G.; Magkonisc, G.; Zekented, K.M. Credit Risk and the Business Cycle: What Do We Know? International Review of Financial Analysis. January 2020, Volume 67. Available online: https://www.sciencedirect.com/science/article/abs/pii/S1057521918307579 (accessed on 6 November 2021).
29. Collin-Dufresne, P.; Goldstein, G.S.; Martin, J.S. The Determinants of Credit Spread Changes. *J. Financ.* **2001**, *56*, 2177–2207. [CrossRef]
30. Fu, X.; Li, M.C.; Molyneux, P. Credit default swap spreads: Market conditions, firm performance, and the impact of the 2007–2009 financial crisis. *Empir. Econ.* **2021**, *60*, 2203–2225. [CrossRef]
31. Hu, K.-H.; Lin, S.-K.; Ching, Y.-K.; Hung, M.-C. Goodness-of-Fit of Logistic Regression of the Default Rate on GDP Growth Rate and on CDX Indices. *Mathematics* **2021**, *9*, 1930. [CrossRef]
32. Moody's Corporation. Annual Default Study: Corporate Default and Recovery Rates, Moody's Investor Service. 2021. Available online: The-performance-of-Moodys-corporate-debt-ratings-Q1-2021-Excel-Supplement-30Apr21.xlsx (accessed on 10 May 2021).
33. IMF. Real GDP Growth. Available online: https://www.imf.org/external/datamapper/NGDP_RPCH@WEO/OEMDC/ADVEC/WEOWORLD (accessed on 10 May 2021).
34. Roch, F.; Roldán, F. Uncertainty Premia, Sovereign Default Risk, and State-Contingent Debt. *IMF Work. Pap.* **2021**, 47. [CrossRef]