Are Default Rate Time Series Stationary? A Practical Approach for Banking Experts*

Gábor Szigel – Boldizsár István Gyűrűs

As the IFRS 9 accounting standard requires banks to recognise impairments based on a forward-looking expected loss concept, banks must estimate the quantitative relationship between default rates and macroeconomic indicators (GDP, unemployment, etc.). In such models, the stationarity of the (usually short) default rate time series is often the most critical issue. In this article, we provide practical advice for banking experts on how (under which circumstances) they can still use short default rate time series in OLS regressions even if those fail regular stationarity tests. We argue that if margin of conservativism is requested for the underlying default rate projections, then applying (seemingly) non-stationary default rate time series in OLS models might not necessarily be problematic.

Journal of Economic Literature (JEL) codes: C22, C53, G21, G32

Keywords: default rates, probability of default, stationarity, time series analysis

1. Introduction

It is not obvious why anyone should care about analysing the stationarity of default rate time series[1] *in general*. After all, if someone wants to use a specific default rate time series in a regression model, they can check the stationarity of it and act accordingly: for example, if they find that the given time series is non-stationary, then they will not use it as a target variable of an OLS regression and if they do not find other alternative modelling approaches, then they can look for other data or simply abandon their publication ambitions in the given topic.

---

* The papers in this issue contain the views of the authors which are not necessarily the same as the official views of the Magyar Nemzeti Bank.

Gábor Szigel: OTP Bank Nyrt., Senior Internal Advisor. E-mail: Gabor.Tamas.Szigel@otpbank.hu
Boldizsár István Gyűrűs: OTP Bank Nyrt., Junior Modeller. E-mail: Boldizsar.Istvan.Gyurus@otpbank.hu

Views expressed in this paper reflect those of the authors and cannot be interpreted as an official opinion of OTP Bank. The authors thank for Mr Atanaz Popov (OTP Bank Nyrt.) his support of their work.

The first version of the Hungarian manuscript was received on 16 June 2023.

DOI: https://doi.org/10.33893/FER.22.4.107

[1] By default rate time series, we mean the default ratios of banks' loan portfolios defined based in the "Basel approach": the default ratio means the share of borrowers in the loan portfolio that will suffer a default event – over 90 days arrears or any other unlikely-to-pay event, e.g. bankruptcy. etc. – within 12 months.

This type of freedom is often not available to banking experts. The IFRS 9 accounting standard and the obligatory supervisory stress testing exercises "force" banks to estimate a quantitative relationship between bank-level risk parameters, e.g. probability of default (PD), or loss given default (LGD), and macroeconomic indicators to calculate the forward-looking adjustment of the former. To meet these external expectations, banks often must be satisfied with less ideal circumstances, such as default rate time series spanning only 10-15 years or (not independently of the aforementioned) contradictory or inconsistent stationarity test results that might change with each annual update.

In this paper, we provide some practical advice to banking experts who are required to work with default rate time series for IFRS 9 or stress testing purposes, where there is no obvious evidence that the series are stationary (or non-stationary).

Our approach is not without precedent as there has been much discussion in the literature on the stationarity of frequently used macroeconomic indicators, such as the GDP time series (examples: *Christiano – Eichenbaum 1990; Rapach 2002; Ozturk – Kalyoncu 2007*). To our knowledge, however, no similar analysis of the general stationarity of default rate time series has been conducted so far.

The article proceeds as follows: Section Two provides some context on the modelling exercises banks face in relation to default rates. Section Three summarises the concept of stationarity. In Section Four, we assess the circumstances under which default rate time series can be regarded as stationary, based on theoretical considerations. In Section Five, we use Monte Carlo simulations to illustrate that short, default rate like stationary time series tend to fail in the most common stationarity (unit root) tests. In Section Six, we show that alternative modelling techniques which can also be used for non-stationary time series often provide less conservative projections. Finally, we present our conclusions.

2. Modelling exercises with default rate time series in banking practice

Banks use quantitative models on the loan portfolio's default rate time series mostly for forecasting purposes. The approach goes usually as follows: first, they quantify the relationship between default rates and other external variables (mostly macroeconomic indicators, such as GDP, unemployment, etc.) based on historical observations and then they create (or take over) future scenarios for macroeconomic variables. Finally, they calculate projections for default rates from those scenarios based on the quantified relationship between them.

Banks prepare such default rate projections under different scenarios for at least four purposes:

- for internal purposes: to support business decisions of the management;
- to estimate expected loss-based impairments under IFRS 9;
- to conduct stress tests to meet supervisory expectations; and
- to estimate long-term (through-the-cycle, or TTC) probability of defaults (PD) (i.e. to calculate their capital requirements according to the IRB[2] approach).

When using the default rate projections for internal purposes, the only expectation of them is that they “work”, i.e. they do not mislead the decisionmakers.

However, in the rest of the cases, external stakeholders such as auditors and supervisors also evaluate those models. These stakeholders have their own expectations: they request banks to find a statistically robust relationship between default rates and macro variables that can be properly documented. For example, during the regular stress testing exercise of the European Banking Authority (EBA), banks are requested to quantify their forecasts for loan losses conditional to very specific macroeconomic scenarios that the European Systemic Risk Board (ESRB) provides (*ESRB 2023*). Under IFRS, banks must estimate expected losses (along with the most important composite: probability of default) using quantitative models[3] for several different macroeconomic scenarios (for the relevant requirements of IFRS 9 models, see also: *Háda* (*2019*), for IRB models, see: *Nagy – Bíró* (*2018*)).

Although it is clear that there is some relationship between the default rates of banks’ loan portfolios and the main indicators describing the general state of the economy, this relationship is usually not robust in time and not independent of the underlying structures; in fact, in this regard one should apply the “Lucas critique”,[4] as was pointed out by *Simons and Rolwes* (*2009*). This was highlighted by the Covid-19 crisis as well: although Hungary suffered an economic contraction in

[2] Internal Ratings-Based

[3] It should be noted though that the IFRS 9 standard and its implementation in EU legislation does not explicitly require the application of quantitative models to calculate forward-looking expected losses. However, in actual practice, auditors and financial supervisory authorities tend to expect the development of such models. Interpreting IFRS 9 literally, an approach in which forward-looking PDs are set to the last observed default rate if no economic recession is expected and are otherwise set to 150 per cent of the previous observation would fulfil the general requirement of using forward-looking indicators, but in reality, auditors and supervisors expect banks to have much more sophisticated models in place.

[4] The Lucas critique – as formulated by *Robert Lucas* (*1976*) – originally states that impacts of changes in economic policy cannot be assessed purely based on dependencies between variables observed in the past and measured by econometric models. The reason is that the underlying structures influencing the decisions of the actors are not constant, and therefore the impacts of policy changes are unpredictable and cannot be reliably prognosticated by models in the long run. Lucas allows that econometric models might provide good forecasts in the short run, but only in cases when the structures and incentives of actors do not change abruptly.

2020 comparable to the Great Recession from 2008, the default rates of banks' loan portfolios remained low and very far away from the high levels of the post-Lehman episode, due to the different characteristics and government responses to this crisis.

Nevertheless, the expectations of auditors and financial supervisors force banks to model the correlations between default rates and macroeconomic indicators based on historical time series and apply those in their forecasts. Although the weaknesses of such models can be adjusted by (expert-based) "overlays", the application of some kind of quantitative model is inevitable. These are mostly simple, OLS-based regression models (although there are alternative approaches, as we show in *Subsection 3.5*.).

3. The concept of stationarity and its role in OLS regressions

In this section, we briefly summarise the concept of stationarity and its necessity in the case of time series applied in OLS models. As there are numerous textbooks on econometrics (e.g. *Békés – Kézdi 2019*) that cover this topic abundantly, we will keep this as short as possible.

3.1. The concept of stationarity

A time series – or to be more precise, the process generating it – is considered stationary if the characteristics of its distribution do not change over time. In banking practice, this definition is often simplified, and a process is considered stationary if its mean and standard deviation are constant over time ("weak stationarity").[5] Time series generated by stationary processes show random fluctuations around a constant mean, from which observations may deviate over the short run, but which they return to sooner or later ("mean reversion").

3.2. What makes processes non-stationary?

Processes and the generated time series can be non-stationary for several reasons:

- *Deterministic trend:* the process follows a time-dependent trend, e.g. it has a constant tendency to increase or decrease. Such processes do not have a constant mean. For example, stock exchange indices are non-stationary, because they tend to increase over the long run.

[5] Note that this is not the exact definition of weak stationarity. More precisely we would require that the mean of the process as well as the autocovariance of any two elements (of any distance away) are constant.

- *Stochastic trend (random walk)*: a process follows a random walk with an autoregressive coefficient between the adjacent observations larger than or equal to 1. For example, individual stock prices can be characterised by random walk, as today's share prices will be always close to yesterday's ones, but it is unpredictable where the prices will be in the long run. Such processes do not have a constant standard deviation (it increases in time) and their mean is not independent of previous realisations of the process,

- *Seasonality:* the mean of the process depends on a seasonal time factor. For example, ice cream sales are seasonal, as people tend to consume more ice cream during the warm summer months.

- *Structural break:* a structural break causes a permanent change of either the mean or other characteristics of the distribution of the process, and therefore they will not be constant over time. For example, sales of candles before and after the invention of electric lighting.

3.3. Stationarity and cyclical variables

We take a closer look at cyclical variables as our subject, as default rate time series are usually also cyclical. At first glance, cyclicality seems to be similar to seasonality: in both cases, there are alternating periods in which our variable takes values characteristic for the given period (winter/summer, crisis/non-crisis). However, there is a big difference: cyclicality – as opposed to seasonality – is not predictable,[6] as the length of the cycles is not constant or known in advance. Therefore, cyclical variables (such as GDP growth) can have constant mean and standard deviation if we look forward over a sufficiently long period. Moreover, the essence of cyclicality is that the process reverts to its mean after swings ("mean reversion"), which is one of the central assumptions of stationarity. Hence, cyclicality does not, *per se*, imply that a certain process is non-stationary.

3.4. Problems with non-stationary variables in OLS regressions: spurious correlations

The stationarity properties of a process are not interesting in isolation, but rather only when the time series generated by the process is used in a regression model. Using non-stationary time series in OLS regressions leads to "spurious correlation", as is well known since at least the 1970s from the article by *Granger and Newbold* (*1974*).

[6] Although in the past the economic literature raised the possibility that the economic cycles obey some kind of "laws of nature" and occur with more or less regular frequency ("the Kondratiev wave"), this idea is not accepted by most economists.

We speak of spurious correlations if variables seem to correlate in the observation period without a real reason,[7] i.e. only due to coincidence. This occurs most likely in the following circumstances:

- when both variables follow a *deterministic trend* (both are permanently increasing): these are the best-known examples of spurious correlations (e.g. the average temperature of Earth and the GDP of Hungary in the last 30 years);
- when both variables follow a *stochastic trend*, e.g. both are random walks. This is less intuitive, but the article of *Granger and Newbold* (*1974*) referenced above showed that spurious correlations can be easily found between random walks as the random shifts might make them change persistently and seemingly in the same direction;
- when both variables are *seasonal* (e.g. ice cream sales, number of drownings);
- when both variables have *structural breaks* or persistent changes (e.g. volume of candle sales and number of work horses in the last 100–150 years: both decreased due to the development of technology).

Spurious correlations (regressions) cannot be used properly for forecasting, because the relationship between the variables that they capture does not exist in reality and therefore will probably also not prevail over the forecast horizon. Typically, such spurious correlations in OLS regressions can be identified by very high R^2 and strongly correlated error terms (with Durbin-Watson statistics far below 1). The mathematical consequence of the autocorrelated error terms is that the t-statistics of regression coefficients (betas) will be overestimated and p-value becomes unreliable. Thus, the statistics based upon which we should decide whether the given explanatory variables have a statistically significant impact on the target variable are useless.

It is important to emphasise though that the decision as to whether our regression is spurious should not be based purely on statistical parameters. Theoretical considerations also need to be taken into account: can a direct causal relationship between the variables be explained, or are they driven by the same root cause, etc.? Only if they are, can one expect the relationship between the variables to be true based on such theoretical arguments.[8]

[7] I.e. without any direct or indirect causal relationship, or at least without a stable general root cause.
[8] Even though the correlation between the variables might be still much weaker in such cases than it is implied by our regression.

3.5. Alternative modelling approaches for non-stationary time series

There are modelling techniques that can capture the relationship between probability of default and macroeconomic variables and yet do not require the stationarity of portfolio-level default rate time series. For example, borrower-level estimation of PD including both borrower characteristics and macroeconomic variables by a logistic regression does not require the use of such time series, and this means there is also no issue with stationarity (as the estimator in such regressions is not OLS). Furthermore, non-stationary variables can also be estimated by OLS if they are co-integrated.

The authors are aware of the existence of such alternative approaches. However, these solutions usually require significantly more data and/or different circumstances that might not be available in the given environment. Therefore, the stationarity of portfolio-level default rate time series remains a relevant topic (at least for a while yet).

4. Stationarity of default rates based on theoretical considerations

4.1. Stationary characteristics of default rate time series

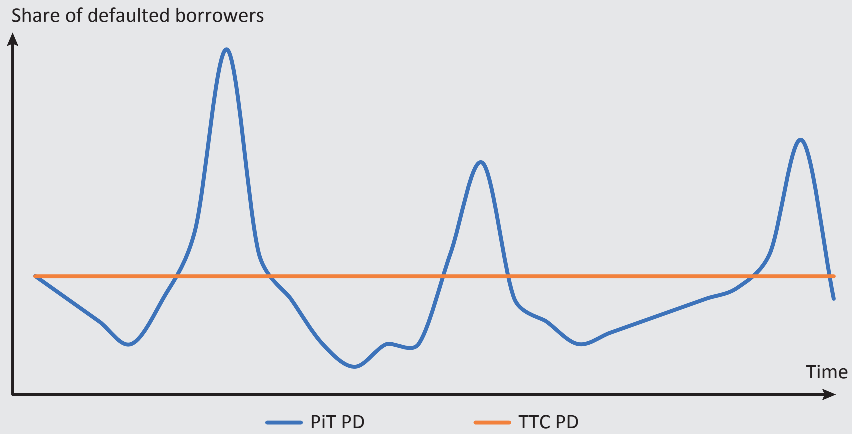
Portfolio-level default rate time series, or to be more precise, the portfolio-level probability of default process generating such, have a number of stationarity properties based on economic theory:

- *Constant mean in the long term:* according to banking practice and the basic concept of the IFRS 9 accounting standard, banks' portfolio-level probability of default fluctuates around a long-term, through-the-cycle average, the so-called TTC PD (*Figure 1*). This would imply that the mean of the process is constant, that is, if the structural characteristics of the underlying loan portfolio (lending standards, composition, external regulation, etc.) are also constant. This is an important limitation which we will return to in *Subsection 4.3*.

- *Mean reversion*: after the high values of a crisis period, default rates tend to revert to lower values as the economic environment improves, thus coming back to a long-term average.

- *Time independency:* as deviations from the TTC PD are driven usually by the economic cycle, in which shocks often result from external and random events (such as the Covid-19 pandemic or the Russian-Ukrainian war), forward-looking realisations of PD do not depend on the actual default rates. Therefore, the cyclicality of PD is not as regular as, for example, a sinus curve (even though it is often depicted like that). On the contrary, it is much more common that "long, calm" periods are followed by a short crisis period (*Figure 1*). Therefore, this cyclicality is not deterministic as opposed to seasonal variables. Although annual default rates are usually autocorrelated (the default rate this year is similar to the one last year), this autocorrelation is not constant and not independent of the actual position of the cycle. It is also important to emphasise that even if there is a strong correlation between adjacent observations of the default rate time series, this does not imply non-stationarity as long as the autocorrelation coefficient is smaller than 1.

Figure 1
Illustration of the TTC and PiT PD



Note: PiT = Point-in-Time (actual), TTC = Through-the-Cycle (long-term average).

All in all, default rate time series are cyclical time series with mean reversion and – under the premises of constant portfolio composition and regulatory environment – a constant mean in the long run. These are stationary characteristics.

Furthermore, there are some general non-stationary characteristics (as presented in *Subsection 3.2*) that can obviously be ruled out in case of portfolio-level PD:

- as PD has a limited value set (between 0–1[9]), it cannot follow a deterministic trend;
- this limited value set also rules out the presence of a stochastic trend;
- PDs are clearly non-seasonal: the Basel regulation defines PD as the probability of default over a 1-year period, and therefore PD cannot exhibit intra-year seasonality.

4.2. Cyclicality and stationarity of default rate time series

Default rate time series are cyclical, but as we pointed out in *Subsection 3.3* cyclicality does not imply non-stationarity. However, another problem can be raised: if we try to regress default rates on other cyclical indicators, we might find relationships that are not causal, but only a consequence of the cyclicality. In banking practice, this problem is usually not relevant, as banks use the "root cause" indicators which represent the economic cycle the best (GDP growth, unemployment, etc.) for modelling purposes (as presented in Section 2).[10]

4.3. Stationarity and structural breaks in default rate time series

There is one factor that can endanger the stationarity of a portfolio-level default rate time series: possible structural breaks. These can be caused specifically by the following phenomena:

- *Changes in definitions or changes in the data collecting process*: these are the most obvious examples of structural breaks, such as the change in the EBA's definition[11] of default in 2017. Although these can be corrected by statistical methods, they inevitably increase modelling uncertainties.
- *Changes in lending conditions:* a bank can change its business strategy voluntarily or as a result of regulatory pressure and enter or withdraw from more risky client segments, which will necessarily change the portfolio's TTC PD (the mean of the process).

[9] In reality, PDs fluctuate usually in an even smaller range: e.g. mortgage loans typically have PDs between 0–5 per cent.

[10] However, the utmost caution is necessary here as there are many macroeconomic indicators that are cyclical, but do not represent the economic cycle the best. Certain components of GDP or their proportion to GDP – despite being often cyclical – might not capture the actual state of the economy the best. For example, exports/GDP might show cyclical patterns, but this indicator is clearly not a good representative for the state of the economy: it can deteriorate both when GDP is increasing or decreasing. Using such indicators to forecast default rates might be very misleading.

[11] Guideline *EBA/GL/2016/07* introduced a new default definition in 2017 that caused a level shift in the default time series of European banks which were not able to implement the new definition retroactively.

- *Changes in the external environment*: there may be other factors independent of banks' own actions which can have an impact on the TTC PD of loan portfolios such as government incentives (increase or decrease in moral hazard that may incentivise borrowers' willingness to pay, such as rescue programmes), changes in the legal environment (e.g. introduction or suspension of the "right of walk away" or bans on evictions, impacts of court decisions) or other cultural factors (e.g. the education process of borrowers).

If such changes are frequent, random and of limited impact, then they will not change a portfolio's TTC PD significantly in the long run, and thus they do not ruin the stationarity of the process. However, if there are only a few, but large structural breaks in the default rate time series, then the underlying process cannot be regarded as having a constant mean and being stationary.

Overall, default rate time series can be considered stationary based on purely theoretical arguments if the presence of significant structural breaks in its historical data can be ruled out with great certainty (or we can at least filter out their effects reliably).

5. Reliability of testing for stationarity in the case of default rate time series

Although, as shown above, it can often be decided based on theoretical-economic arguments whether a specific default rate time series can be assumed to be stationary, formal stationarity tests must be also performed, both due to practical and documentational reasons. The most widespread formal approaches to test stationarity tests are unit root tests such as the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, as well as some others. However, as highlighted in the following sections by a short literature review and by our own Monte Carlo simulations, these tests are not reliable on *short* time series and can often lead to a false identification of stationary processes as non-stationary.

This is a problem, because default rate time series – especially in the Central Eastern European region – tend to be short. As banks only started to collect default data from about the mid-2000s in line with the introduction of Basel II, the data collected by the 2020s ranges back 15 years at best. Additionally, this time span covers only one whole economic cycle (the post-Lehman crisis episode and the subsequent recovery, as data from the Covid-19 pandemic cannot be used for modelling due to the special circumstances of that period).

5.1. Challenges with testing on short time series

It is well-known in the literature that stationarity tests (in reality: unit root tests) are unreliable on short time series. *Arltová and Fedorová* (*2016*) showed via their Monte Carlo simulations that most stationarity tests provide misleading results with a sample size less of than 50 observations and show stationary processes to be non-stationary. The KPSS test, which – in contrast to other stationarity tests – operates with a null hypothesis that the process is stationary, provided the most precise results.[12] Autocorrelation in the underlying process makes the situation even worse: an autocorrelation of above 0.5 drastically reduces the reliability of all but the KPSS test, while an autocorrelation of above 0.7 will make even the KPSS test unreliable. This poses a problem for our particular case, as empirical experience – to which we will come back in *Section 5 and 6* – suggests that default rate time series (on annual frequency) tend to be highly autocorrelated (on a magnitude of 0.5–0.7).

There are alternative solutions to test stationarity on small samples: some authors provide alternative critical values for short time series, e.g. *Jönsson* (*2006*) for the KPSS test, and *Otero and Smith* (*2003*) and *Cheung and Lai* (1995) for the ADF test. Using these decreases the probability that the test will falsely identify a stationary process as non-stationary; in return, however, it increases the probability of the other type of error (identifying a non-stationary process as stationary). Furthermore, application of these alternative critical values has not become widespread. *Cochrane* (*1991*) also points out that highly autocorrelated processes – such as probability of defaults generating the default rates – might always be misrecognised by the tests on finite samples. All in all, the "philosophers' stone" of stationarity testing for small samples could not be found, and this is not surprising given that the basic challenge results from the fact that there are just simply not enough observations in small samples to see whether the underlying process has a constant mean and other stable distribution characteristics.

The way the length of the time series impacts the results of stationarity tests is illustrated by *Figure 2*: here, we see the default rates of the Moody's Ba[13] corporate rating class as observed each year between 1920 and 2006,[14] and the relevant t-statistics of the ADF test for each year calculated from the beginning of the time

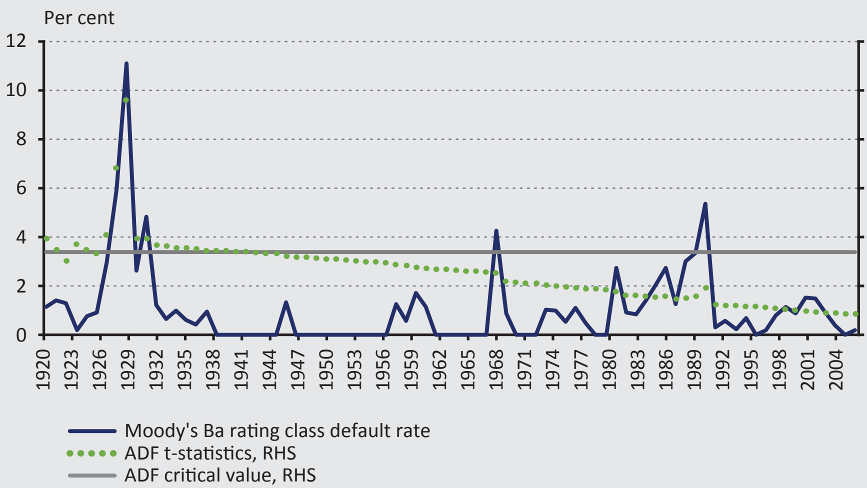
[12] While most stationarity tests (e.g. the ADF test) use the null-hypothesis that the time series has a unit root and thus the alternative hypothesis is that the data are stationary, the null-hypothesis of the KPSS test is the opposite, namely that the time series is stationary.

[13] We selected the Ba rating class because, in contrast to better rating classes, there have been enough default rate observations differing from zero over the years. Such default rates are much more similar to those of banks' loan portfolios, that include also more vulnerable borrowers than the AAA type of corporations.

[14] *Moody's* (*2007:20–21*): *Exhibit 21 – Annual Issuer-Weighted Corporate Default Rates by Letter Rating, 1920–2006*.

series up to the given year.[15] If the t-statistic is below the critical value of –2.6, the ADF test identifies the process as stationary at a confidence level of 10 per cent. As can be seen, for the first 20 observations/years, the t-statistic is larger than the critical value, implying that the process is non-stationary. This judgement of the test could also be confirmed by visual inspection to some extent: one can easily see an increasing trend in the default rates up to the 1930s. Over the long run, however, the ADF test clearly identifies the Moody's default rates to be stationary at the 10 per cent confidence level (moreover, from the 1970s at the 1 per cent confidence level as well). This example underlines how much the length of the observation period impacts the results of the test.

Figure 2
Default rate time series of Moody's Ba corporate rating class between 1920 and 2006, and the corresponding t-statistics and critical values of the ADF test



Note: t-statistic at time t shows what the t statistic of the ADF test would be if we calculated it for a time span from 1920 to time t. For more information on the number of lags applied in the test: see Footnote 15. Critical values are calculated without trend, but with constant for T=25 at the 10 per cent confidence level. Although critical values change slightly as the number of observations change, the difference is so small that it would not have an impact.

Source: Moody's (2007) and own calculations

[15] Calculated with constant. When choosing the appropriate lag order, we did the following: using the full-length time series, we assessed which lags were significant at a confidence level of 10 per cent and we omitted insignificant lags from the regression of the ADF test. As no lags were significant in the case of this particular time series, we did not include any of them – thus running a simple Dickey–Fuller test instead of the ADF. We note that including more lags would not change the decreasing trend of the t-statistic, though it would cause a shift in its value upwards. The ADF test would identify the Ba default rates as stationary up to 6 lags on the full length of the data up to 2006.

5.2. Results of our Monte Carlo simulations

The unreliability of stationarity tests on small samples is well-known from the literature, but the relevant research does not focus on default rate like processes, but rather on ones based on normal distribution. Default rate time series tend to have special types of distributions: their value set is limited (portfolio-level default rates cannot be outside of the 0–100 per cent range) and they are cyclical (crisis/ non-crisis episodes following each other). Therefore, portfolio-level default rates cannot follow normal distributions.

This motivated us to assess the reliability of the ADF and KPSS tests for default rate like processes with the use of Monte Carlo simulations. We introduce two novelties compared to other authors such as *Arltová and Fedorová* (*2016*):

- We use the ADF and KPSS tests at the same time for the generated time series to assess whether this would increase the overall reliability of the tests ("at least one would recognise stationary processes correctly"). We pick these two tests because they are the most widespread ones in banking practice, and, even more importantly, they have an opposite null hypothesis and can thus complement each other (see also: Footnote 12).
- We test explicitly default rate like processes – ones with limited value sets and cyclical properties.

We assess the stationarity of eight processes, altogether: four non default rate like and four default rate like.

The *non default rate like processes* – which we test as a "control group" and to replicate the results of the literature – are the following:

Process No. 1: Stationary, non-autocorrelated, following normal distribution;

Process No. 2: Random walk (non-stationary);

Process No. 3: Stationary, highly autocorrelated (ρ = 0.7);

Process No. 4: Stationary, very highly autocorrelated (ρ = 0.9).

The following four are the *default rate like processes*. These processes were created in a way that they are all stationary:

Process No. 5: Limited value set, non-autocorrelated, cyclical: a variable with a randomly changing crisis and non-crisis mode (with two constant means, one for crisis, one for non-crisis), limited to the 0–1 value set;

Process No. 6: Limited value set, autocorrelated, mean-reverting cyclical: autoregressive process with random crisis shocks and with an additional error correction, that pulls back the process to its long-term average based on how high the deviation from this average was during the previous realisation. Limited to the 0–1 value set;

Process No. 7: Generated from corporate bankruptcy rates in Hungary by bootstrapping, non-autocorrelated: we chose randomly from the realised corporate bankruptcy rates in Hungary as presented in *Subsection 6.1* in more detail. The results will be default rate like, non-autocorrelated cyclical time series ranging between 0–1*;*

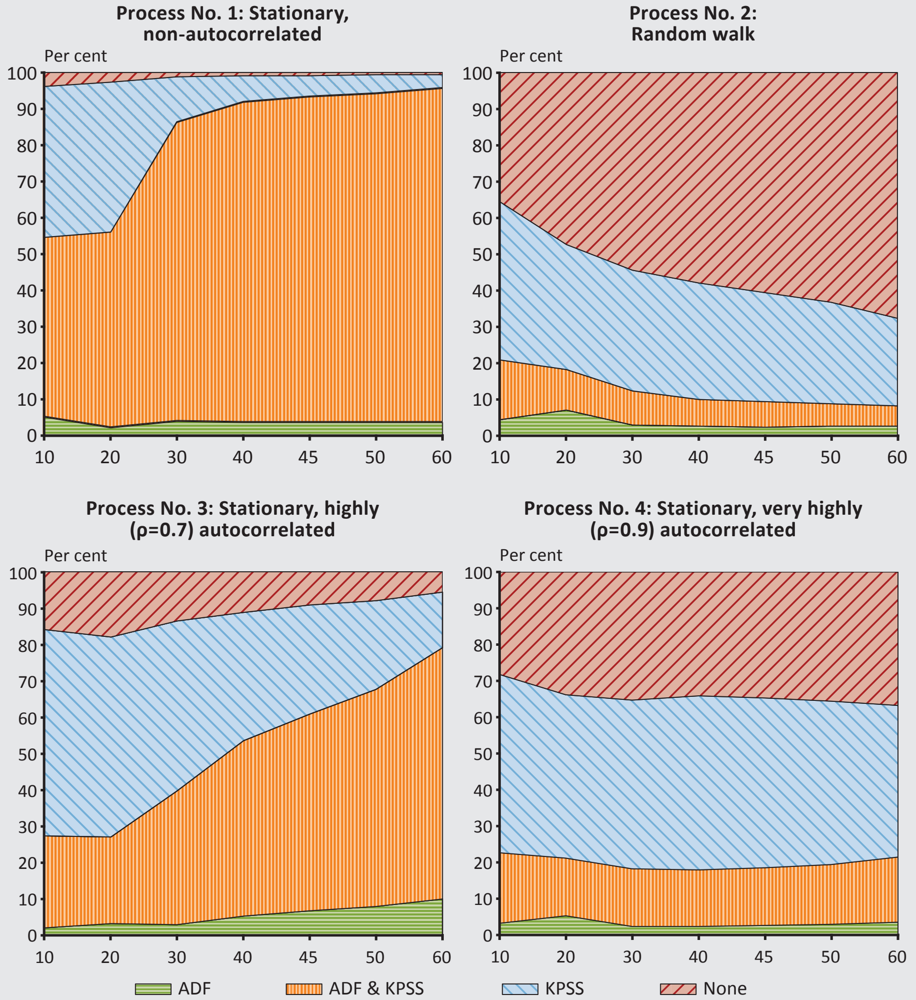
Process No. 8: Generated by fixed-length sections of the Moody's Ba rating class's historical default rate, autocorrelated: we chose fixed-length sections of the real Moody's Ba default rate time series from randomly selected starting date. The benefit of this approach is that the $\rho = 0.5$ autocorrelation of the Moody's default rates will be kept in the simulated time series.

The exact definition of the above processes can be found in the *Appendix*. We generated 100,000 simulations for each process (except for *Process No. 8*, because the length of the original Moody's time series limited our possibilities; for further details, see the *Appendix*). The length of the generated time series varied between 10 and 60 observations (10, 20, 30, 40, 45, 50, 60). After this the relevant test statistics of the ADF and KPSS test were calculated using an embedded Python package (called "statsmodels.tsa.stattools") and compared to the critical values at the 10 per cent confidence level from *Dickey – Fuller* (*1979*) and *Kwiatkowski et al.* (*1992*). Finally, the generated time series were categorised based on their test results:

- stationary according to only the ADF,
- stationary according to only the KPSS,
- stationary according to both,
- non-stationary according to both.

The simulation results for the non default rate like processes are summarised by *Figure 3*. These are in line with the literature: the ADF test often provides "false negative" results (recognising stationary time processes as non-stationary) for short samples consisting of 10–20 observations. Although the KPSS test mostly identifies the stationary processes correctly (i.e. as stationary) it fails to correctly categorise the non-stationary random walk process (Process No. 2), and therefore it is also unreliable. Both tests tend to accept their null hypothesis for the very short samples irrespective of the real characteristics of the underlying process. In the case of strongly autocorrelated stationary processes (Process No. 3–4), the performance of both tests deteriorates further for the very short samples compared to how they behaved for the non-autocorrelated processes. In the case of lower autocorrelation, the precision of the tests improves with the sample size, but in the case of very high autocorrelation (ρ = 0.9) even increasing the sample size does not help.

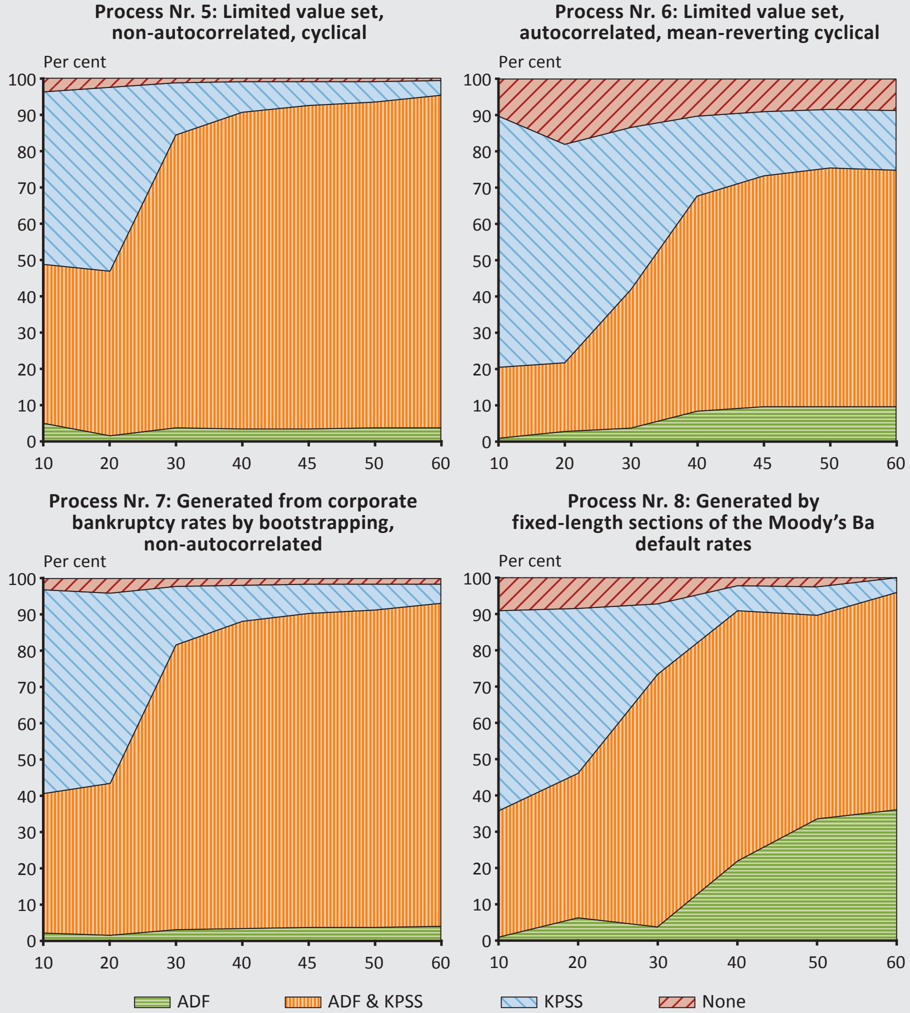
Figure 3
Results of the ADF and KPSS tests for the non default rate like processes for different time series length



Note: The figure shows what percentage of the simulated time series of different length (10, 20, 30, 40, 45, 50, 60) "passes" the stationarity tests (are recognised as stationary by the tests). The labels are the following: "ADF" = stationary only according to ADF, "ADF&KPSS" = stationary according to both tests, "KPSS" = stationary only according to KPSS, "None" = non-stationary according to both tests.

Figure 4 shows similar results for the default rate like (known to be stationary) processes as for the non default rate like stationary processes. Results for the non-autocorrelated processes (Process No. 5 and 7) are very much similar to the ones for the non-autocorrelated normal distribution process (Process No. 1): ADF often provides a "false negative" assessment for 10–20 observations and then its precision improves with an increase in sample size. For the autocorrelated processes

Figure 4
Results of the ADF and KPSS tests for the default rate like processes for different time series length



Note: The figure shows what percentage of the simulated time series of different length (10, 20, 30, 40, 45, 50, 60) "passes" the stationarity tests (are recognised as stationary by the tests). The labels are the following: "ADF" = stationary only according to ADF, "ADF&KPSS" = stationary according to both tests, "KPSS" = stationary only according to KPSS, "None" = non-stationary according to both tests.

(No. 6 and 8), ADF performs even worse on very small samples (T=10 or 20), and its reliability improves only gradually with more observations (it often misrecognises the stationary processes even with a sample size of 30 or 40). Although the KPSS test identifies the default rate like processes correctly as stationary with a high probability, the merits of this result are limited due to the known weakness of the test to misrecognise non-stationary time series as stationary. Even worse, KPSS fails to accept the stationarity of the autocorrelated Process No. 6 and 8. in a relatively high proportion of the cases.

All in all, the ADF test is not appropriate for correctly identifying stationary processes on very small samples (10–20 observations) and the same can be said for the KPSS test: even though the latter can recognise a stationary process as stationary if it is not highly autocorrelated, it will do so with non-stationary processes as well and furthermore it fails to correctly identify stationarity if the underlying process – as default rate time series are – is highly autocorrelated.

Therefore, if a time series consisting of only 10–20 observations fails these stationarity tests, that still does not mean it is certain that the underlying process is indeed non-stationary.

5.3. Increasing the sample by using a higher frequency of observations

The previous section also showed that the reliability of stationarity tests – especially that of the ADF test – improves as the sample size increases. In banking practice, it is a widespread approach to "enlarge" the sample sizes of default rate time series by increasing the observation frequency from annual to quarterly, i.e. by taking default rates for Q1 2007, Q2 2007, etc. instead of for 2007, 2008, etc. (whereby the default rate still reflects the share of defaulting borrowers within 12 months, only the frequency of starting points increases). With this approach, modellers can significantly increase the length of time series, as we have, for instance, 52 observations from the period between 2007–2019 at quarterly frequency as opposed to just 13 observations on an annual basis.

Nonetheless, increasing data frequency is not a "silver bullet" to solve stationarity issues with short time series: *Pierse – Snell* (*1995*) argue[16] that it is less effective than increasing the length of the sampling period. Additionally, there is a special challenge with increased frequency in the case of default rate time series: as the default ratio must be measured for a 12-month observation period (and

[16] The underlying argument is that "mean reversion" prevails only with the increase of the length of the observation period, but not with the increase of frequency. The same article also shows that in the case of a sufficiently long time series, data frequency does not have an impact on the results of stationarity (unit root) tests.

measurement for a shorter period is practically not feasible[17]), adjacent default rate observations at quarterly frequency will overlap by 75 per cent and will therefore be even more autocorrelated. This is a problem, because – as shown in Subsection 5.2 – the more autocorrelated the time series are, the less reliably the stationarity tests will work on them (especially KPSS).

Empirical experience shows that default rates on an annual frequency are highly autocorrelated in themselves. Autocorrelation increases further with the switch to quarterly frequency due to the overlap: for example, the corporate bankruptcy rates presented in Section 6 show an autocorrelation of 0.76 at annual frequency, but that jumps to 0.98 at quarterly frequency! Practical experience of the authors suggests that the autocorrelation of default rates of banks' different portfolios ranges between 0.6–0.8 at annual frequency, but 0.9–0.99 at quarterly frequency. If a process is so highly autocorrelated as given by the latter range, then even a sample size of 50–60 is not sufficiently large for the stationarity tests (especially for KPSS) to work reliably.

All in all, if we have annual default rate data from only 15–20 years, then transforming them into quarterly frequency will not solve the reliability problems of the stationarity tests: what we gain from the increase in the sample size, we might easily loose through the additional autocorrelation we cause.

6. Challenges with alternative modelling approaches

The false identification of stationary default rate time series as non-stationary takes a toll if it pushes banks to choose alternative modelling approaches that might generate worse forecasts for loan portfolios' PD than model types that require stationarity. What banking modellers often do to be able to use OLS estimations (which requires stationarity) is the following: they difference the (seemingly) non-stationary default rate time series and, as the differences of the adjacent default rate observations tend to constitute a stationary time series, they then enter those as the target variable into the OLS regressions. This practice is very widespread with many examples in the literature (*Balatoni – Pitz 2012*; *Gál 2019*), and it is often completely justified and correct.

[17] A series of challenges can be raised: intra-year seasonality of defaults, or the fact, that the default trigger of the 90+ days arrears can often simply not occur within a quarter in case of newly originated loans.

However, differencing the variables may also be unjustified as the models for the differenced target variable provide unreasonable results. In such cases, we can speak of over-differencing. *Cochrane* (*2018*) describes the underlying mechanism: the model with the differenced variable can often show only a much weaker relationship between the target and the explanatory variables, because such models do not only try to quantify the correlation (and causal relationship, ideally) between the variables, but implicitly also the timely co-movement between them (or in other words: the impulse response of the target variable on the explanatory variable).

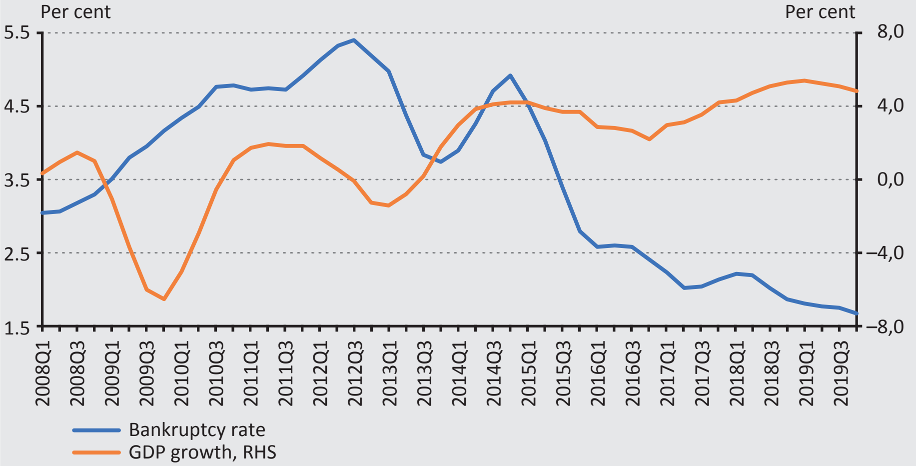
Translating this for our default rate case: PD models applied on the differences of the portfolio-level default rates will show a much weaker relationship between those and the macroeconomic variables, ultimately leading to forecasts in which the predicted default rates are completely insensitive to the state of the external environment. This is especially a problem, if estimates are expected by third parties (supervisors, auditors) to be conservative – similarly to the case of stress testing. In the following subsection, we illustrate a case of over-differencing with an example.

6.1. Example of the challenges of over-differencing

In our example, we regress the 12-month bankruptcy rate of Hungarian companies – which is a publicly available default rate like variable – on the 12-month GDP growth rate of Hungary, at a quarterly frequency. In order to simulate the data availability that is typical in Hungarian banks, the bankruptcy rate time series used is limited to the period between Q1 2008 and Q4 2019 (this is the time span for which Hungarian banks usually have default rate data to model with).[18] Overall, we have data from only one complete economic cycle. *Figure 5* provides an overview of the two time series we are using.

[18] In reality, corporate bankruptcy rates are available for the period between Q1 1996 – Q4 2022. We omitted the data between 2020 and 2022, i.e. the Covid-19 period from our modelling exercises, because that was a very special crisis period: GDP dropped abruptly and significantly, and then rebounded in a similar way, while bankruptcy and default ratios hardly increased, due to government and regulatory interventions (e.g. payment moratorium on loans). Treating this period appropriately still poses a challenge in banks' PD modelling exercises as well.

Figure 5
Corporate bankruptcy rate and real GDP growth in Hungary between 2008–2019



Source: MNB, Macroprudential indicators

Looking at the figure above, even without conducting any stationarity test one can immediately suspect that the bankruptcy rate will be stationary, because it shows a downward trend. It also seems likely in advance that an OLS regression will find a relationship between the bankruptcy rate and GDP growth as the former takes high values during the 2008–2013 crisis episode and low values during the recovery, while the latter has values the other way around. However, this relationship will be dominated by the experiences from the post-2008 crisis episode which will limit its usefulness for forecasting purposes – as the next crisis might not be like that of the previous one (the Covid-19 crisis was in fact not similar).

Conducting the ADF test on the bankruptcy rate finds it indeed non-stationary (t-statistic: –1.04, presence of a unit root can be rejected at a probability of 73 per cent). The first difference of the bankruptcy rate is clearly stationary by the test though (t-statistic: –2.03, presence of a unit root can be rejected at a probability of 96 per cent).[19] (We also note that GDP growth appears to be clearly stationary by the ADF test even for the period between 2008–2019, and for the full length of the time series).

[19] For both tests, the maximum number of lags was 7.

Then, we prepare two different models: we regress both the level and the difference of the bankruptcy rate[20] on GDP growth. The results are shown in *Table 1*. GDP is a significant explanatory variable in both models (at a 10 per cent confidence level), but the coefficients – expressing the strength of the relationship – differ by a magnitude: the β is 0.2 in the model on the level-variable, but is only 0.02 in the case of the model on the difference. Of course, none of the models is particularly good (R^2 is low, although the F-statistic is significant, but the Durbin-Watson statistics are very low, especially in the model on the level variable). However, our motivation here was not to find the best possible model, but rather to illustrate the consequences of treating non-stationarity problems by differencing. Moreover, in banking practice, similar situations regarding the availability of data and the quality of the models are common.

Table 1
Results of the two OLS regressions

Number of model	Target variable	Explanatory variables	Constant	β (of GDP growth)	R^2
1	Bankruptcy rate (level)	Constant, GDP growth	0.04***	–0.21***	0.27
2	Bankruptcy rate (difference)	Constant, GDP growth	Not significant	–0.02*	0.06

*Note: We also tried to use the difference of GDP growth as an explanatory variable in the models, but it turned out to be insignificant. Asterisks next to the coefficients have the usual meanings: *** significant at 1 per cent, * significant at 10 per cent.*

In model-technical terms, the model on differences seems to be the correct one, as the OLS estimations are run on a stationary target variable. However, this model shows hardly any relationship between GDP growth and the bankruptcy rates: according to this, an economic contraction of –1 per cent would lead to an increase of barely 0.02 percentage point in the bankruptcy rate in a quarter (+0.08 percentage points in a year).

[20] We note that it would be more appropriate to use some kind of logistically transformed version of the target variables in both models so that forecasts based on these models cannot fall outside the reasonable value set of default rates between 0 and 1. However, this would render the interpretation of the model coefficients more complicated, and therefore we follow this simplified approach for the purposes of illustration.

The insensitivity of the model on difference is reflected if we use the two models to predict bankruptcy rates during the Covid-19 crisis and compare those forecasts to what actually happened. To do so, we assume "perfect foresight", i.e. that the modellers would have known the realised GDP trajectories in advance. The results of this comparison are shown in *Table 2*. As we can see, none of the models would have been particularly successful in predicting bankruptcies: the model on levels would have provided strongly fluctuating forecasts for the first two years of the Covid-19 pandemic and would have overestimated the actual bankruptcy rates by far (even for the year of the economic rebound). It would have hit the bankruptcy rate in 2022 though. The model on differences would have predicted a relatively flat bankruptcy rate despite the big economic ups and downs and would have underestimated the rates over the whole forecast horizon, but especially in 2022. It is true that the model on differences would have been closer to reality – but only because bankruptcy rates proved to be just as insensitive to the deterioration of the economic circumstances in the first year of the pandemic as the model on differences.

Table 2
Predictions of the two models and reality

Quarter	Bankruptcy rate as predicted by the models one year in advance (%)		Bankruptcy rate in reality (%)	GDP growth (%)
	Model on levels	Model on differences		
2019Q4	–	–	1.7	4.9
2020Q4	4.9	1.7	1.4	–4.5
2021Q4	2.4	1.3	1.7	7.2
2022Q4	3.0	1.1	2.9	4.6

Note: In case of the model on difference, the forecast is prepared on the basis of the actual bankruptcy rate each year, to which we added the change of PD based on the forecasted GDP growth as predicted by the model assuming "perfect foresight".

However, we can assess the performance of the two models not only based on their precision in forecasting reality, but also based on their ability to meet the expectations of banks' external stakeholder – such as supervisors or auditors – with a preference for *conservative* approaches. For example, at the onset of the pandemic, when expected GDP growth was –4.5 per cent under perfect foresight, the model on differences would have predicted a flat bankruptcy rate, whereas the model on level would have anticipated a substantial rise in the number of bankruptcies similar to the levels experienced after the Lehman collapse. We have good reasons to assume that stakeholders targeting conservative, "prepare-for-the-bad-times" approaches (supervisors, auditors) would have hesitated to accept the results of the model on differences and would have preferred the model on levels (even though reality would have proven them wrong). It is also hard to imagine that European banks could submit projected default rate trajectories like those coming from the model on differences, for example during an EBA stress testing exercise (which uses assumptions on GDP growth similarly severe to the first year of the Covid-19 crisis).

All in all, it would have been a mistake to dismiss the model on levels only because the level of bankruptcy rates fails on stationarity tests, because:

- it would have been probable that the test is wrong on the stationarity of the process generating the data because of the small sample size, and that bankruptcy rates are stationary in reality (although we did not assess the possibility of a structural break here);
- economic theory suggests that there is a causal relationship between economic growth and bankruptcy rates; and
- the alternative of the model on levels would have resulted forecasts that are completely insensitive to the change in the economic environment.

At the same time, it is also important to emphasise that the model on levels could have been used for predictions only if its limitations were presented transparently during the interpretation of results: namely, that it exclusively reflects the experience of the Great Recession after 2008, which might not be relevant for the actual economic environment (as it proved not to be relevant, in fact).

7. Conclusions and practical advice

In this paper, we analysed whether loan portfolios' default rate time series, or to be more precise the underlying processes generating them, can be considered stationary based on theoretical-economic arguments and whether the relevant unit root tests are appropriate to identify them in practice, with due consideration of the usually short length of these time series (observations often only from 15–20 years).

We found that default rate time series can be assumed to be stationary if there are no structural breaks in them, i.e. the underlying default definition, data gathering process, lending standards, legal environment and other borrower incentives, etc. did not cause a level-shift in the expected value of probability of default.

While to some extent these assumptions hold in many cases, default rate time series that are collected from an observation period with a relatively short time span often fail the regular stationarity tests (unit root tests). This is not surprising, as stationarity tests are unreliable for short samples (less than 30 observations), especially if the assessed time series (and the underlying data generating process) are highly autocorrelated. As default rate time series tend to have these properties, it is probable that stationarity tests identify them falsely as non-stationary, especially if the number of observations is less than 30.

Using non-stationary variables in an OLS regression is not advisable as it can lead to "spurious correlations". Accordingly, modellers choose alternative modelling approaches such as differencing of the target variable if it is thought to be non-stationary. However, this is often a bad choice for banks' default rates, at least for purposes which require conservativism (such as stress testing or IFRS 9 forward-looking expected loss models): although models on differences usually have better properties in econometric terms, they tend to quantify weak relationships between default rates of loan portfolios and the macroeconomic environment. Predictions based on such models will often be unsensitive to the expected deterioration of economic circumstances and hence not conservative enough.

As a conclusion of our analysis, we provide the following advice to banking experts:

- Take extra care when assessing the stationarity of default rate time series if the data is of less than 20-30 years (be careful in general, of course).
- As stationarity (unit root) tests with opposite null hypotheses (such as ADF and KPSS) tend to accept their own null hypothesis on small samples, it is a good rule of thumb to accept their results if those are consistent (meaning that at least one of the tests with opposing hypotheses rejected its null).
- Short default rate time series should not be assessed as non-stationary solely based on the results of stationarity tests, especially if they fail only one of the tests. Expert judgement should be used to consider theoretical-economic arguments that might justify the stationarity of the default rate time series.
- When undertaking the above, special attention should be paid to the presence of structural breaks: if these are absent, then short default rate time series can be assumed to be stationary even if they fail stationarity tests or if different tests provide inconsistent results.
- If the purpose of default rate modelling requires conservative predictions (e.g. in the case of stress testing), then it is not necessarily incorrect to use the level of default rates as target variables in OLS regressions, even if they fail stationarity tests. Especially if the only alternative is a model on differences that is insensitive to the explanatory variables (to the change of the external environment)
- However, when interpreting model results, it should be presented transparently in what way the above model-technical choices might distort the predictions, and when (under what circumstances) they might work well and when they cannot.

References

Arltová, M. – Fedorová, D. (2016): *Selection of Unit Root Test on the Basis of Time Series Length and Value of AR(1) Parameter*. Statistika, 96(3): 47–64. https://doaj.org/article/ca92920816ef46e0b9e874642271805e

Balatoni, A. – Pitz, M. (2012): *A működőtőke hatása a bruttó nemzeti jövedelemre Magyarországon* (*The effect of direct investment on Hungary s gross national income*). Közgazdasági Szemle (Economic Review), 59(1): 1–30. https://www.kszemle.hu/tartalom/cikk.php?id=1284

Békés, G. – Kézdi, G. (2021): *Data Analysis for Business, Economics and Policy*. Cambridge University Press. https://doi.org/10.1017/9781108591102

Cheung, Y.-W. – Lai, K.S. (1995): *Lag Order and Critical Values of the Augmented Dickey-Fuller Test*. Journal of Business & Economic Statistics, 13(3): 277–280. https://doi.org/10.2307/1392187

Christiano, L.J. – Eichenbaum, M. (1990): *Unit roots in real GNP: Do we know, and do we care?* Carnegie-Rochester Conference Series on Public Policy, 32, 7–61. https://doi.org/10.1016/0167-2231(90)90021-C

Cochrane, J.H. (1991): *A critique of the application of unit root tests*. Journal of Economic Dynamics and Control, 15(2): 275–284. https://doi.org/10.1016/0165-1889(91)90013-Q

Cochrane, J.H. (2018): *A Brief Parable of Over-Differencing*. February 11. https://www.johnhcochrane.com/research-all/a-brief-parable-of-overdifferencingnbsp. Downloaded: 29 May 2023.

Dickey, D.A. – Fuller, W.A. (1979): *Distribution of the Estimators for Autoregressive Time Series with a Unit Root*. Journal of the American Statistical Association, 74(366a): 427–431. https://doi.org/10.1080/01621459.1979.10482531

ESRB (2023): *Macro-financial scenario for the 2023 EU-wide banking sector stress test*. European Systemic Risk Board, 31 January 2023. https://www.esrb.europa.eu/mppa/stress/shared/pdf/esrb.stress_test230131~c4980ac646.en.pdf. Downloaded: 29 May 2023.

Gál, Z. (2019): *Az FDI szerepe a gazdasági növekedés és a beruházások területi differenciálódásában Magyarországon* (*The role of FDI in Hungary's economic growth and the territorial differentiation of investments*). Közgazdasági Szemle (Economic Review), 67(6): 653–686. https://doi.org/10.18414/KSZ.2019.6.653

Granger, C.W.J. – Newbold, P. (1974): *Spurious regressions in Econometrics*. Journal of Econometrics, 2(2): 111–120. https://doi.org/10.1016/0304-4076(74)90034-7

Háda, A. (2019): *Banking Supervisors Tracing the Transition to IFRS 9*. Financial and Economic Review, 18(4): 77–111. https://doi.org/10.33893/FER.18.4.77111

Jönsson, K. (2006): *Testing Stationarity in Small and Medium-Sized Samples when Disturbances are Serially Correlated*. Working Paper No. 20, Department of Economics, Lund University. https://lup.lub.lu.se/search/files/5598406/2061568. Downloaded: 24 August 2023.

Kwiatkowski, D. – Phillips, P.C.B. – Schmidt, P. – Shin, Y. (1992): *Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?* Journal of Econometrics, 54(1–3): 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

Lucas, R.E. (1976): *Econometric policy evaluation: A critique*. Carnegie-Rochester Conference Series on Public Policy, 1, pp. 19-46. https://doi.org/10.1016/S0167-2231(76)80003-6

Moody's (2007): *Corporate Default and Recovery Rates, 1920-2006*. Special comment, Moody's Investors Service – Global Credit Research, February. https://www.moodys.com/sites/products/DefaultResearch/2006400000429618.pdf. Downloaded: 11 September 2023.

Nagy, G. – Bíró, G. (2018): *PIT and TTC problems related to IRB PD parameter estimation in the light of supervisory reviews*. Economy and Finance, 5(3): 250–278.

Otero, J. – Smith, J. (2003): *The KPSS Test with Outliers*. Computational Economics. 29, 59–67. https://doi.org/10.1007/s10614-005-9008-0

Ozturk, I. – Kalyoncu, H. (2007): *Is Per Capita Real GDP Stationary in the OECD Countries? Evidence from a Panel Unit Root Test*. Ekonomski Pregled, 58(11): 680–688. https://mpra.ub.uni-muenchen.de/9635/

Pierse, R.G. – Snell, A.J. (1995): *Temporal aggregation and the power of tests for a unit root*. Journal of Econometrics, 65(2): 333–345. https://doi.org/10.1016/0304-4076(93)01589-E

Rapach, D.E. (2002): *Are Real GDP Levels Nonstationary? Evidence from Panel Data Tests*. Southern Economic Journal, 68(3): 473–495. https://doi.org/10.1002/j.2325-8012.2002.tb00433.x

Simons, D. – Rolwes, F. (2009): *Macroeconomic Default Modeling and Stress Testing*. International Journal of Central Banking, 5(3): 177–204. https://www.ijcb.org/journal/ijcb09q3a6.pdf

Appendix: Formal definition of the processes used in the Monte Carlo simulations

We used the following processes in our Monte Carlo simulation exercise:

1. Stationary non-autocorrelated $(\rho = 0)$ *process*

$$y_t \sim \mathcal{N}(\mu;\ \sigma^2)$$

2. Random walk, non-stationary unit root $(\rho = 1)$ *process*

$$y_t = y_{t-1} + \varepsilon_t$$

$$y_0 \sim \mathcal{N}(\mu; \sigma^2)$$

$$\varepsilon_t \sim \mathcal{N}(0;\ \sigma^2)$$

3–4. Stationary, strongly-correlated $(\rho = 0.7 \text{ and } \rho = 0.9)$ *processes*

$$y_{t+1} = \rho y_t + \varepsilon_t$$

$$y_0 \sim \mathcal{N}(\mu;\ \sigma)$$

$$\varepsilon_t \sim \mathcal{N}\left((1-\rho)\mu; \sqrt{1-\rho^2}\sigma\right)$$

where $\mathcal{N}()$ stands for the normal distribution.

5. Bounded, non-autocorrelated, cyclic stationary process

$$y_t = \begin{cases} \mathcal{N}(\mu * 0.8;\ \sigma), & Rnd_t > \ p \\ \mathcal{N}(\mu * 2.4;\ \sigma), & Rnd_t \leq \ p \end{cases},$$

where the generated values are bounded to the 0.003–1 region,

$$y_0 \sim \mathcal{N}(\mu;\ \sigma),$$

where $\mathcal{N}()$ stands for the normal distribution, Rnd_t is a random value between 0 and 1 corresponding to time t, and *p* is the probability of crisis shock (chosen to be 12.5 per cent). There are two states in this process: the crisis phase, when the generated values fluctuate around a "crisis mean value" and the non-crisis phase, when the generated values fluctuate around a "non-crisis mean value". The crisis and non-crisis mean values and the value for the probability of crisis shock was chosen in such a way that the overall mean of the process is μ, while the variance is σ. 5 per cent was selected for the value of μ (which is not an unrealistic estimate for the default rate of a loan portfolio), and thus the boundary associated with the 0.003–1 interval rarely applied. Note that the choice of the minimum (0.3%) is set to be non-zero, because of the non-zero PD in the Basel regulation.

6. Bounded, autocorrelated, mean-reverting stationary

$$Shock\_d_t = \begin{cases} 0, \ Rnd_t > p \\ 1, \ Rnd_t \le p \end{cases},$$

$$y_t = \begin{cases} 3 * y_{t-1} + \mathcal{N}(0;\ \sigma) + (y_{t-1} - \mu * 0.8)/3, Shock\_d_t - Shock\_d_{t-1} = 1 \\ y_{t-1} + \mathcal{N}(0;\ \sigma) + (y_{t-1} - \mu * 0.8)/3, Shock\_d_t - Shock\_d_{t-1} \ne 1 \end{cases},$$

wehere *Shock_d* denotes a shock dummy and the generated values are limited to the 0.003–1 region.

$$y_0 \sim \mathcal{N}(\mu;\ \sigma),$$

where $\mathcal{N}()$ stands for the normal distribution, Rnd_t is a random value between 0 and 1 corresponding to time t, and *p* is the probability of crisis shock (chosen to be 12.5 per cent). *The effects of the crisis shock are somewhat different compared to those in the previous processes.* If a crisis shock occurs in a non-crisis year the following datapoint will be three times that of the previous one. If the crisis shock occurs in a crisis year, the process will simply take the value of the previous observation as a basis. These values are then supplemented by a normal error term and a mean reversion term, the latter of which is responsible for pulling the value back the current value of the series to the average. This reversion becomes stronger as the values start to diverge. 5 per cent was selected for the value of μ and thus the boundary associated with the 0.003–1 interval rarely had to apply. Note that the choice of the minimum (0.3%) is set to be non-zero, because of the PD floor in the Basel regulation.

Although it might seem like this process is non-stationary, after careful numerical analysis (e.g. Monte Carlo simulation), one can quickly deduce that the mean and the variance are independent of time.

7. Generated from corporate bankruptcy rates in Hungary by bootstrapping, non-autocorrelated process[21]

Table 3 summarises the corporate bankruptcy rates in Hungary between 1996 and 2022 (i.e. 27 observations). During this process, selecting the data points is done by randomly selecting values from this table, allowing for repetition.

Table 3
Corporate bankruptcy rates in Hungary between 1996–2022

Year	Bankruptcy rate (%)	Year	Bankruptcy rate (%)	Year	Bankruptcy rate (%)
1996	2.9	**2005**	2.8	**2014**	4.9
1997	2.5	**2006**	3.1	**2015**	2.8
1998	2.3	**2007**	3.0	**2016**	2.4
1999	2.2	**2008**	3.3	**2017**	2.1
2000	2.5	**2009**	4.2	**2018**	1.9
2001	2.7	**2010**	4.8	**2019**	1.7
2002	2.6	**2011**	4.9	**2020**	1.4
2003	3.0	**2012**	5.2	**2021**	1.7
2004	2.8	**2013**	3.8	**2022**	2.9

Source: MNB (2023) Report on Financial Stability, May 2023, Macroprudential indicators, Chart M11. (https://www.mnb.hu/letoltes/melleklet-abrai-2023-majus.xlsx)

8. Generated by fixed-length sections of the Moody's Ba rating class's historical default rate, autocorrelated

One takes the Moody's Ba rating class's historic default rate (shown in *Figure 2*) and selects fixed sized sections of it (10, 20, 30, 40, 50, 60 observations) using every starting point. As the longest interval is 60 observations long, there are only 27 possible starting points in the 87-datapoint long dataset. To preserve statistical consistency, it was decided to use 27 randomly selected sections for all the other lengths (10, 20, 30, 40, 50) as well.

[21] Note that a slightly different version of this process was also examined, in which the technique of bootstrapping was used on the Moody's Ba rating class's historical default rate instead of the Hungarian corporate bankruptcy rates. The results of this version were almost identical, and so it is not presented here.