



**Faculty of
Mathematics
and Informatics**

VILNIUS UNIVERSITY
FACULTY OF MATHEMATICS AND INFORMATICS
MODELLING AND DATA ANALYSIS
MASTER'S STUDY PROGRAMME

MEASURING THE QUALITY OF THE THROUGH-THE-CYCLE PROBABILITY OF DEFAULT PARAMETERS ESTIMATION

Master's thesis

Author: Aistė Žečkytė

VU email address: aiste.zeckyte@mif.stud.vu.lt

Supervisor: Doc., Dr. Dmitrij Celov

Vilnius

2023

Contents

1	Introduction	4
2	Theoretical context	5
2.1	Basel regulations	5
2.2	PIT and TTC PDs	6
3	Literature review	7
4	Methodology	9
4.1	Asymptotic Single Risk Factor model	9
4.2	Hybrid rating model	11
4.3	Estimation of the parameters ρ , B and α	11
4.3.1	General Method of Moments	12
4.3.2	Two-step regression	13
4.4	Credit cycles	13
4.5	Monte Carlo method	14
4.6	Gaussian Mixture Model	15
4.7	Controlled simulation experiment	16
4.7.1	Scenarios and assumptions	16
4.7.2	Algorithm of the experiment	17
5	Empirical results	19
5.1	Normality of simulated credit cycles	19
5.2	Parameter estimates	21
5.2.1	Scenario 1	21
5.2.2	Scenario 2	23
5.2.3	Hybrid estimation method	24
6	Conclusions	25
7	Appendix A. Figures and tables	29
8	Appendix B. R code	50

Measuring the Quality of the Through-the-Cycle Probability of Default Parameters Estimation

Abstract

The purpose of this thesis is to investigate the impact of non-normally distributed components on the estimation of the through-the-cycle probability of default (TTC PD) parameters in the Asymptotic Single Risk Factor (ASRF) model, which is used by financial institutions to calculate capital requirements under the Basel III regime. The study uses a controlled Monte Carlo simulation experiment involving 10,000 simulations for two countries, Lithuania and Germany, associated credit cycles. Each simulation involves generating a credit cycle, creating portfolios with specific parameters, and evaluating the parameter estimates for each portfolio using the General Method of Moments (GMM), a two-step regression (TSR), and a combination of them. The results show that deviations from normality in the credit cycle significantly affect the accuracy of parameter estimates, with larger deviations resulting in greater errors. In the more realistic scenario, the quality of the portfolio also plays a role in the accuracy of the estimates. The study also finds that the GMM most of the time performs better than TSR in estimating the TTC PD parameters and in certain cases, the most accurate results are provided by a combination of these two methods. Additionally, the results show that neither the GMM nor the TSR method is suitable for low default portfolios.

Key words: probability of default (PD), through-the-cycle (TTC), point-in-time (PIT), Gaussian Mixture Model, Asymptotic Single Risk Factor model, General Method of Moments, two-step regression.

Ekonominio ciklo įsipareigojimų nevykdymo tikimybės parametų vertinimo kokybės matavimas

Santrauka

Šio darbo tikslas – ištirti nenormaliai pasiskirsčiusių komponentų įtaką ekonominio ciklo įsipareigojimų nevykdymo tikimybės parametų įvertinimui asimptotinio vieno rizikos faktoriaus modelyje, kurį pagal Bazelio III reikalavimą finansų institucijos naudoja kapitalo poreikiui apskaičiuoti. Tyrime naudojamas kontroliuojamas Monte Karlo simuliacijų eksperimentas, apimantis 10 000 simuliacijų dviejų šalių kredito ciklams – Lietuvai ir Vokietijai. Kiekviena simuliacija apima kredito ciklo generavimą, portfelių su konkrečiais parametrais generavimą ir kiekvieno portfelio parametų įverčių vertinimą naudojant bendrąjį momentų metodą, dviejų žingsnių regresiją ir šių dviejų metodų derinį. Rezultatai rodo, kad kredito ciklo nukrypimai nuo normalumo reikšmingai veikia parametų įverčių tikslumą, o didesni nuokrypiai lemia didesnes paklaidas. Realistiškesniame scenarijuje portfelio kokybė taip pat turi įtakos įvertinimų tikslumui. Tyrime taip pat nustatyta, kad bendrasis momentų metodas dažniausiai veikia geriau nei dviejų žingsnių regresija, įvertinant ekonominio ciklo įsipareigojimų nevykdymo tikimybės parametrus, o tam tikrais atvejais tiksliausias rezultatus pateikia šių dviejų metodų derinys. Be to, rezultatai rodo, kad nei bendrasis momentų metodas, nei dviejų žingsnių regresija nėra tinkami žemo nemokumo portfeliams.

Raktiniai žodžiai: ekonominio ciklo įsipareigojimų nevykdymo tikimybė, laiko momento įsipareigojimų nevykdymo tikimybė, Gauso mišinio modelis, asimptotinio vieno rizikos faktoriaus modelis, bendrasis momentų metodas, dviejų žingsnių regresija.

Abbreviations

- ASRF - *Asymptotic Single Risk Factor*
- BCBS - *Basel Committee on Banking Supervision*
- EU - *European Union*
- GMM - *General Method of Moments*
- IRB - *internal rating based*
- MCMC - *Markov Chain Monte Carlo*
- ODF - *observed default frequency*
- OLS - *ordinary least squares*
- PD - *probability of default.*
- PIT PD - *point-in-time probability of default*
- RWA - *risk-weighted assets*
- TSR - *two-step regression*
- TTC PD - *through-the-cycle probability of default*
- UCM - *unobserved components model*

1 Introduction

The financial industry has used risk classification systems for a long time. According to Basel II and the subsequent European Union (EU) regulations, commercial banks must determine the appropriate level of risk-weighted assets (RWA) that will remain stable despite changes in the credit cycle. The Basel risk-weight formula is used to calculate the capital needed to cover unexpected losses. The formula is based on the probability of default (PD), loss given default, and exposure at default. The regulation requires the use of “long-run average one-year default rates” for the PD which is the main research object of the current thesis. If the probability of default is not through-the-cycle (TTC), it will vary with the economic cycle, and the capital requirements will also change. It is even more necessary to have a consistent methodology for using TTC probability of default in RWA formulas and to clearly understand and quantify the state of the economic cycle and probability of default under the Basel III regime.

Many recent papers (Carlehed and Petrov [2], Rubtsov and Petrov [15], Oeyen and Celis [8]) have argued that both through-the-cycle and point-in-time (PIT) probability of default have their own advantages and disadvantages and that a good steering system should ideally utilize both concepts. Because most models in practice are hybrid models, it is important to transform hybrid PDs into TTC PDs using a point-in-time-ness (PITness) parameter as an additional parameter in the Asymptotic Single Risk Factor (ASRF) model. Financial institutions are required to follow the latter model for capital calculations, where past and future fluctuations are expected to be explained by both systematic and idiosyncratic movements. Systematic movements reflect default risk due to a common factor, while idiosyncratic movements are specific to a particular credit risk profile.

ASRF model assumes that both, idiosyncratic and common components are Gaussian. However, in real-life applications, when applied to relatively small heterogeneous observations, the data could not meet such assumptions. For this reason, the purpose of the thesis is to investigate the quality of the TTC PDs parameters estimation by conducting a controlled Monte Carlo simulation experiment.

In order to examine how non-normally distributed components affect estimates of the parameters, two different scenarios are introduced:

1. Common component is generated by the stochastic simulations according to the structural unobserved components model (UCM) and this component is taken as given. The idiosyncratic component is generated as a normal random variable.
2. Common component is the same as in 1. The idiosyncratic component is generated as a random variable from the mixture of Gaussian distributions.

A controlled simulation experiment was conducted using 10,000 Monte Carlo simulations. Each simulation involved generating a credit cycle, creating several portfolios of different sizes with specific parameters, and evaluating the parameter estimates for each portfolio using the General Method of Moments (GMM), two-step regressions (TSR) and a combination of those two. For the experiment, generating credit cycles, two countries have been chosen: Lithuania and Germany. Since ASRF model

assumes normality of the common credit cycles related components, thus, in each simulation, the normality hypothesis for the common component is tested.

The obtained results showed that in the cases of both countries, common components, i.e. credit cycle, deviations from the normal distribution are not random. In the case of the first scenario, it was observed that larger deviations from normality affect the accuracy of parameter estimates. The results of the second scenario are quite similar, but here the accuracy of the estimates depended more on the quality of the portfolio (i.e. on TTC PD). Additionally, the results highlighted the low default portfolio problem that is quite often encountered in reality.

Content-wise, the thesis is structured as follows: Section 2 introduces some context about the Basel regulations and main definitions. A concise review of the literature is presented in Section 3. Sections 4 and 5 describe all the methodology and empirical results, respectively. Finally, conclusions and possible proposals for further work development are presented in Section 6. The computations for the thesis were conducted using R version 4.1.2.

2 Theoretical context

2.1 Basel regulations

Basel regulations, also known as the Basel Accords, are a set of international banking standards that were developed by the Basel Committee on Banking Supervision (BCBS) to promote the stability and integrity of the financial system [11]. These regulations are primarily concerned with the risk management practices of banks, including the measurement of credit risk, operational risk, and market risk.

Basel regulations specify the minimum requirements for banks to use internal models for estimating the probability of default (PD) [13]. These requirements include the use of appropriate data, assumptions, and methodologies, as well as the implementation of robust risk management practices to ensure the accuracy and reliability of the estimates.

One key aspect of these regulations is the requirement for banks to use a through-the-cycle approach to PD estimation, which means that the estimates should be based on a long-term perspective and take into account the potential impact of cyclical and structural changes in the economic environment [14]. This approach is intended to provide a more realistic and conservative estimate of credit risk, as compared to a point-in-time approach that only considers the current economic conditions.

Basel I was the first set of international banking standards developed by the BCBS [9]. It was introduced in 1988 and focused on the measurement of credit risk. Under Basel I, banks were required to hold a minimum amount of capital based on the credit risk of their assets. The minimum capital requirement was calculated using a standardized approach, which did not take into account the specific risk characteristics of individual banks. While Basel I was seen as a step forward in promoting the stability and soundness of the financial system, it had some limitations, including its lack of differentiation between different types of credit risk and its reliance on simple risk weights.

Basel II was introduced in 2004 as an update to Basel I, with the aim of improving the risk sensitivity

of the capital framework and addressing some of the limitations of the previous framework [10]. It introduced three “pillars” to the capital framework: a minimum capital requirement, a supervisory review process, and market discipline. Under Basel II, banks were given the option to use internal models to estimate their capital requirements, subject to regulatory approval. The internal models approach (also known as an internal rating based or IRB) was intended to be more risk-sensitive and allow banks to better reflect the specific risk characteristics of their portfolios. However, the implementation of Basel II was criticized for being too complex and for creating incentives for banks to optimize their capital ratios through the use of complex financial instruments.

Finally, Basel III was introduced in 2010 in response to the global financial crisis, with the aim of strengthening the resilience of the financial system by increasing the quality and quantity of banks’ capital and liquidity [12]. It introduced additional requirements for the quality and quantity of capital, as well as new liquidity ratios to ensure that banks have sufficient liquid assets to meet their short-term funding needs. Basel III also made some changes to the IRB approach of Basel II, including the introduction of a through-the-cycle approach to credit risk measurement, which takes into account the potential impact of cyclical and structural changes in the economic environment. Basel III is currently being implemented by banks around the world, and its implementation is being closely monitored by regulators to ensure that it is effective in promoting the stability and soundness of the financial system.

2.2 PIT and TTC PDs

In this thesis, we understand the term “PIT PD for an obligor” to mean the probability that the obligor will default within the next twelve months. It is obvious, that the point-in-time PD is conditional on the information that we have on time t , i.e. on time when this probability is estimated. Theoretically, this information could be split into two components: the state of the common factors at time t , and other information that is possessed about the obligor at time t [2]. We will use a single asymptotic common risk factor Z , also referred to as systematic risk.

Definition 2.1. *The point-in-time probability of default (PIT PD), denoted by $p_i(z)$ for an obligor i , is the probability that the obligor will default within the next twelve months if the common factor Z has the value z , all else being equal.*

By definition 2.1 we mean that “other information” about the obligor is constant and only systematic variable Z varies.

Definition 2.2. *The trough-the-time probability of default (TTC PD), denoted by q_i for an obligor i , is the average PIT PD. Here the average is over the common factor Z , i.e. over all states of the economic.*

Here, again, the TTC PD q_i depends only on information known at time t . Similar definitions can be introduced for the portfolio level.

Definition 2.3. *The point-in-time probability of default (PIT PD), denoted by $p_P(z)$ for a portfolio P , is the average of the PIT PDs for the obligors in the portfolio if the common factor Z has the value z ,*

all else being equal:

$$p_P(z) = \frac{1}{\#P} \sum_{i \in P} p_i(z). \quad (2.1)$$

Here $\#P$ is the number of the obligors in the portfolio.

Definition 2.4. *The through-the-cycle probability of default (TTC PD), denoted by q_P for a portfolio P , is the average of the TTC PDs for the obligors in the portfolio:*

$$q_P = \frac{1}{\#P} \sum_{i \in P} q_i. \quad (2.2)$$

3 Literature review

In many frameworks, it is considered that IRB modelled PDs have a hybrid nature: they are neither perfectly point-in-time nor through-the-cycle (Aguais *et al* [1], Carlehed and Petrov [2], Rubtsov and Petrov [15], Oeyen and Celis [8]). To comply with Basel regulations, banks must have estimates of both PIT and TTC probabilities of default. As a result, various methods for obtaining these estimates have been proposed by different authors.

One of the sophisticated approaches is proposed by Aguais *et al* [1]. The authors first categorize obligors into different sectors and then gather default data for each sector. Then, default statistics are translated to Merton's distance to default (described more precisely in Section 4) by using inverse standard normal cumulative distribution function and averaged across all sectors to obtain the final estimates. It results in credit indexes (common components) Z for each sector which describe the level of default risk within economic sectors. A modelled PD is commonly estimated by a logistic regression model including deviations of the mentioned credit indexes from their long-run equilibrium levels as additional explanatory factors. Since authors do not define the degree of PITness, their method is only fit for 100% PIT models, i.e. not hybrid PDs as commonly modelled in IRB. After all, by setting credit index deviations to zero and keeping the estimated regression coefficients unchanged, authors derive to the assumed TTC PD.

The next available methodology is suggested by Carlehed and Petrov [2]. Similarly to Aguais *et al*, they begin with an assumption of the hybrid PD availability, which is then decomposed to pure PIT and TTC PDs. The authors assume that the observed default rate follows the asymptotic single risk factor Merton-Vasicek model, in accordance with Basel regulations. Carlehed and Petrov assume that there is a global portfolio for each given segment, with credit conditions that impact all obligors uniformly through a common cycle Z and Vasicek correlation ρ . The authors use the General Method of Moments (GMM) to estimate two parameters: B , which corresponds to the long-term default frequency for a sector, and ρ , which is the correlation between the average portfolio and the economic cycle. The degree of PITness is estimated with additional assumptions which say that the sub-portfolio is formed of the obligors with a stable risk profile.

PIT and TTC theory in Oeyen and Celis [8] paper is fundamentally the same as in Carlehed and Petrov, however, their estimation procedure differs. The authors use linear regression to calculate the

TTC PD by calibrating the Vasicek correlation ρ and the PITness parameter α . This methodology simplifies the statistical approach for calibrating TTC PD and reduces data requirements. Their TTC PD is time-varying as they allow for changes in the idiosyncratic risk of any given obligor. However, the time-dependent intercept causes problems with ordinary least squares (OLS) and direct implementation would give poor correlation estimates. Oeyen and Celis circumvent this by taking the first-order derivative. The authors note that Vasicek correlation must be estimated first before determining TTC PD. This is done using internal observed default frequency (ODF) and the common factor Z for each (sub)sector. Oeyen and Celis also suggest a top-down, heuristic segmentation framework in order to provide justification for any modeller-based definitions of the common factor.

Entirely different to Carlehed-Petrov and Oeyen-Celis is Rubtsov and Petrov [15] framework. PIT-TTC transformation is done based on rating-grade PDs instead of obligor-specific PDs. The authors assume that the TTC PD for each segment is stochastic and use logistic regression to estimate the boundaries of the rating grades based on the quantiles of the empirical obligor distribution. They use logit scores to track the movement of the economic cycle over time and adjust the empirical quantiles accordingly. This is done by analyzing the migrations of non-defaulted customers from one time period to the next. It allows to effectively remove the influence of the business cycle on the rating boundaries. Thus, the degree of PITness is determined by the change in observed portfolio level PD for behavioural customers belonging to each rating bucket.

Table 1 summarizes the results of these papers in few dimensions (see Table 1), excluding Aguais *et al* that omits explicit assessment of the level of PITness.

Table 1: Comparison of TTC methods.

Dimension	Carlehed & Petrov (2012)	Rubtsov & Petrov (2016)	Oeyen & Celis (2019)
Underlying interaction between TTC and PIT PDs	TTC means an unconditional PD, PIT means a PD conditional on the state of the economic credit cycle	TTC PD is rating-grade specific, PIT PD is then calculated in the following calibration step	TTC means an unconditional PD, PIT means a PD conditional on the state of the economic credit cycle
Dependence on an external credit cycle	Explicit	Implicit	Explicit
Estimating the level of PITness	GMM	Logistic scoring, and relative (dynamic) portfolio movements	Linear regression

Based on the comparison provided in Table 1, in the following work, it was chosen to examine two directly competing methodologies with an explicit link to the external credit cycle, i.e. GMM introduced by Carlehed and Petrov [2] and two-step regression introduced by Oeyen and Celis [8].

4 Methodology

4.1 Asymptotic Single Risk Factor model

Vasicek, in his seminal 2002 paper [19], has shown that Merton's model can be extended to a specific Asymptotic Single Risk Factor credit portfolio model. It is assumed that the obligor defaults on his credit obligations when the obligor cannot meet its liabilities because the value of the obligor's assets at the loan maturity T is lower than the contractual value B of its payable obligations. Let A_i denotes the value of the i -th obligor's assets. Dynamics through the time of A_i is described by the process:

$$dA_i = \mu_i A_i dt + \sigma_i A_i dW_i(t). \quad (4.1)$$

Here μ_i denotes the asset-specific drift, σ_i represents volatility and $W_i(t)$ denotes a Wiener process.

The solution for equation (4.1) can be derived by first dividing both sides by A_i and using Itô's formula [7]. Thus, the value of the assets at T can be expressed as:

$$\log A_i(T) = \log A_i + (\mu_i - \frac{1}{2}\sigma_i^2)T + \sigma_i\sqrt{T}X_i. \quad (4.2)$$

Here X_i is a random variable that follows the standard normal distribution.

The probability of default for the i -th obligor at maturity time T is the probability of the value of the obligor's assets A_i being lower than the threshold value B_i :

$$q_i = \mathbb{P}(A_i(T) < D_i) = \mathbb{P}(X_i < B_i) = \Phi(B_i), \quad (4.3)$$

where

$$B_i = \frac{\log D_i - \log A_i - \mu_i T + \frac{1}{2}\sigma_i^2 T}{\sigma_i\sqrt{T}} \quad (4.4)$$

and Φ is the cumulative normal distribution function.

Now, consider a portfolio that consists of n loans – all in equal amounts. Let the probability of default on every loan be p and assume that, for any two companies, the values of assets of those companies are correlated with a coefficient ρ . In addition, assume that all loans have the same maturity term T . Let L_i be the gross loss on the i -th obligor's loan so that $L_i = 1$ if the obligor defaults and 0 otherwise. In more mathematical form we can express gross loss as:

$$L_i = \mathbb{1}_{\{X_i < B_i\}} = \begin{cases} 1, & \text{when } X_i < B_i, \\ 0 & \text{otherwise.} \end{cases} \quad (4.5)$$

The portfolio percentage gross loss then can be expressed by

$$L = \frac{1}{n} \sum_{i=1}^n L_i. \quad (4.6)$$

By the central limit theorem, if the events of the obligor's defaults in the portfolio were independent of each other and identically distributed random variables, the portfolio loss distribution would converge to a normal distribution as the size of the portfolio increases. However, the events of defaults are not independent, thus the conditions of the central limit theorem are not satisfied and L is not asymptotically normally distributed. Nevertheless, it turns out, that the distribution of portfolio loss converges to a limiting form as derived in Vasicek [19].

In particular, as mentioned earlier in this section, asset values of the borrowers A_i are pairwise correlated with a coefficient ρ for any two companies. It follows that variables X_i in equation (4.2) have pairwise correlation ρ too. Because of it, X_i can be expressed with jointly standard normal variables Z and ε_i and represented as:

$$X_i = \sqrt{\rho}Z + \sqrt{1 - \rho} \varepsilon_i. \quad (4.7)$$

Here $Z, \varepsilon_1, \dots, \varepsilon_n$ are mutually independent random variables distributed according to the standard normal distribution. The variable Z is systematic, or in other words, a common factor of the portfolio that affects all the loans in the portfolio and can be interpreted as an economic index or the state of the economy. It follows then, that the term $\sqrt{\rho}Z$ is the obligor's exposure to the common factor and the term $\sqrt{1 - \rho} \varepsilon_i$ accounts for obligor-specific risk.

The probability of the portfolio loss can be evaluated as the expectation over the common factor Z of the conditional probability given Z . First of all, fixing the common factor Z , the conditional probability of default for any one obligor in the homogeneous portfolio is:

$$p_i(Z) = \mathbb{P}(L_i = 1|Z) = \mathbb{P}(X_i < B_i|Z) = \Phi\left(\frac{\Phi^{-1}(q_i) - \sqrt{\rho}Z}{\sqrt{1 - \rho}}\right). \quad (4.8)$$

This provides the default probability of the i -th obligor under given conditions, i.e. given the state of the economy. Averaging over the common factor Z , which represents various states of the economic cycle, we obtain the unconditional (or TTC) default probability that makes the primary focus of the current thesis:

$$q_i = \mathbb{E}_Z[p_i] = \Phi(B_i). \quad (4.9)$$

By inverting (4.8), for each obligor in the portfolio we can calculate:

$$B_i = \sqrt{\rho}Z + \sqrt{1 - \rho}\Phi^{-1}(p_i) \quad (4.10)$$

and then the final result is:

$$q_i = \Phi\left[\sqrt{\rho}Z + \sqrt{1 - \rho}\Phi^{-1}(p_i)\right]. \quad (4.11)$$

Denoting B as the B_i of the average obligor or central tendency obligor, the conditional probability of the portfolio loss can be represented by:

$$p_P(Z) = \Phi\left(\frac{B - \sqrt{\rho}Z}{\sqrt{1 - \rho}}\right) \quad (4.12)$$

and as above

$$q_P = \Phi(B) = \Phi \left[\sqrt{\rho}Z + \sqrt{1-\rho}\Phi^{-1}(p_P) \right]. \quad (4.13)$$

4.2 Hybrid rating model

In the subsection above, we described two boundary rating philosophies, so-called PIT and TTC PDs, and each of which has its advantages and disadvantages [2]. In real modelling applications, most models vary with the cycle to a certain extent, but not fully, i.e. they are so-called hybrid PD models, which are neither perfectly point-in-time nor through-the-cycle [1]. In this context, the degree of PITness should be introduced.

Let α be a real number between 0 and 1, i.e. $0 \leq \alpha \leq 1$. This variable will be called the degree of PITness of the rating model. The PITness parameter is an additional parameter to the ASRF model, which expresses the dependency of a modelled hybrid PD on a common factor Z [8]. Hybrid PD model for an obligor i can be considered as in (4.8), but with correlation parameter ρ replaced with $\tilde{\rho} = \rho\alpha^2$:

$$p_{i,\alpha}(Z) = \Phi \left(\frac{B_i - \sqrt{\rho}\alpha Z}{\sqrt{1 - \rho\alpha^2}} \right). \quad (4.14)$$

When $\alpha = 1$, it is truly PIT PD, and when $\alpha = 0$, it is truly TTC PD model. Likewise, hybrid PD for portfolio can be simply derived from (4.12) in the same way replacing ρ with $\tilde{\rho} = \rho\alpha^2$:

$$p_{P,\alpha}(Z) = \Phi \left(\frac{B - \sqrt{\rho}\alpha Z}{\sqrt{1 - \rho\alpha^2}} \right). \quad (4.15)$$

Mimicking the steps in the previous subsection, we get:

$$q_i = \Phi \left[\sqrt{\rho}\alpha Z + \sqrt{1 - \rho\alpha^2}\Phi^{-1}(p_{i,\alpha}) \right] \quad (4.16)$$

and analogously

$$q_P = \Phi \left[\sqrt{\rho}\alpha Z + \sqrt{1 - \rho\alpha^2}\Phi^{-1}(p_{P,\alpha}) \right]. \quad (4.17)$$

4.3 Estimation of the parameters ρ , B and α

As mentioned in the literature review, parameters will be estimated using two different methods: the General Method of Moments and two-step regression. In the following sections, we will describe those methods in more detail. The important aspect of the derivation of formulas is that we consider observed default frequency (ODF) to be analogous to PIT PD as in [8]. This simplification is useful for the estimation of Vasicek correlation ρ , the parameter B , and the PITness parameter α .

4.3.1 General Method of Moments

Taking the mean and the variance of (4.12) and using the simplification that is mentioned above, we get:

$$\begin{aligned} E[\Phi^{-1}(ODF_t)] &= E\left[\frac{B - \sqrt{\rho}Z_t}{\sqrt{1-\rho}}\right] \\ &= \frac{1}{\sqrt{1-\rho}} (E[B] - \sqrt{\rho}E[Z_t]) \\ &= \frac{B}{\sqrt{1-\rho}} \end{aligned} \quad (4.18)$$

and

$$\begin{aligned} V[\Phi^{-1}(ODF_t)] &= E[(\Phi^{-1}(ODF_t))^2] - E[\Phi^{-1}(ODF_t)]^2 \\ &= \frac{B^2 + \rho}{1-\rho} - \left(\frac{B}{\sqrt{1-\rho}}\right)^2 \\ &= \frac{\rho}{1-\rho}. \end{aligned} \quad (4.19)$$

Then, we can define m as the mean and σ as the standard deviation of the series $\Phi^{-1}(ODF_t)$, i.e.:

$$m = \frac{B}{\sqrt{1-\rho}} \quad \text{and} \quad \sigma = \sqrt{\frac{\rho}{1-\rho}}. \quad (4.20)$$

Identifying moments, and solving for ρ and B we get

$$\hat{\rho} \approx \frac{\hat{\sigma}^2}{1 + \hat{\sigma}^2} \quad \text{and} \quad \hat{B} \approx \frac{\hat{m}}{\sqrt{1 + \hat{\sigma}^2}}. \quad (4.21)$$

For α parameter estimation Carlehed and Petrov introduce the term ‘‘portfolio with stable risk profile’’ [2] which is mentioned in the assumptions valid in this work (subsection 4.7). Conditional on a portfolio with a stable risk profile, one way to estimate PITness parameter is to use (4.14). Considering two-time points that represent a severe change in credit conditions, i.e. reflect the extent of boom-burst cycle similar to global financial crisis of 2008 or coronavirus outbreak. Let without loss of generality Z_1 denotes the local peak in credit cycle and Z_2 denotes the subsequent trough point, then, for each obligor i :

$$\Phi^{-1}(p_{i,\alpha}(Z_1)) = \frac{B_i - \sqrt{\rho}\alpha Z_1}{\sqrt{1 - \rho\alpha^2}} \quad (4.22)$$

and

$$\Phi^{-1}(p_{i,\alpha}(Z_2)) = \frac{B_i - \sqrt{\rho}\alpha Z_2}{\sqrt{1 - \rho\alpha^2}}. \quad (4.23)$$

Averaging over the obligors (portfolio) and taking the difference between (4.23) and (4.22), we get:

$$\Delta \left(\frac{1}{\#P} \sum_{i \in P} \Phi^{-1}(p_{i,\alpha}(Z)) \right) = \frac{-\sqrt{\rho}\alpha \Delta Z}{\sqrt{1 - \rho\alpha^2}}, \quad (4.24)$$

where the difference in the points of the corresponding functions is denoted as Δ . In (4.24) average of

B_i is cancelled out as it was assumed that the risk profile is stable over time. From this equation the degree of PITness α can be solved:

$$\hat{\alpha} = \frac{\Delta \left(\frac{1}{\#P} \sum_{i \in P} \Phi^{-1}(p_{i,\alpha}(Z)) \right)}{\sqrt{\hat{\rho} \left[(\Delta Z)^2 + \left(\Delta \left(\frac{1}{\#P} \sum_{i \in P} \Phi^{-1}(p_{i,\alpha}(Z)) \right) \right)^2 \right]}}. \quad (4.25)$$

Since the estimate of Vasicek correlation $\hat{\rho}$ depends on ODF moments, theoretically, there is no endogeneity problem assessing the level of PITness by GMM approach.

4.3.2 Two-step regression

Two-step regression is another parameters estimation approach suggested and applied by Oeyen and Celis [8]. Likewise in GMM approach, in this framework, Vasicek correlation ρ must be estimated before determining PITness parameter α . To do that, first of all, we rely on OLS regression to estimate the coefficients β_0 and β_1 in:

$$\begin{aligned} \Phi^{-1}(ODF_t) &= \frac{B - \sqrt{\rho} Z_t}{\sqrt{1 - \rho}} \\ &= \frac{B}{\sqrt{1 - \rho}} - \frac{\sqrt{\rho}}{\sqrt{1 - \rho}} Z_t \\ &= \beta_0 + \beta_1 Z_t. \end{aligned} \quad (4.26)$$

Equating the regression coefficients gives

$$\hat{\rho} \approx \frac{\hat{\beta}_1^2}{1 + \hat{\beta}_1^2} \quad \text{and} \quad \hat{B} \approx \hat{\beta}_0 \sqrt{1 - \hat{\rho}}. \quad (4.27)$$

When ρ is already estimated, (4.15) readily implies

$$\begin{aligned} \Phi^{-1}(p_{P,\alpha}(Z_t)) &= \frac{B - \sqrt{\rho} \alpha Z_t}{\sqrt{1 - \rho \alpha^2}} \\ &= \frac{B}{\sqrt{1 - \rho \alpha^2}} - \frac{\sqrt{\rho} \alpha}{\sqrt{1 - \rho \alpha^2}} Z_t \\ &= \gamma_0 + \gamma_1 Z_t. \end{aligned} \quad (4.28)$$

Much like in (4.26), OLS can be employed to estimate coefficients γ_0 and γ_1 , from which approximation for PITness parameter α is derived:

$$\hat{\alpha} \approx \frac{\hat{\gamma}_1}{\sqrt{(1 + \hat{\gamma}_1^2 \hat{\rho})}}. \quad (4.29)$$

4.4 Credit cycles

Credit cycles, which will be used as a common factor, can be simulated according to the framework described by Celov and Comunale [3]. This section will present the structural unobserved components

model (UCM) used for stochastic simulations. A great advantage of such simulation is that we have full knowledge of simulated elements and, hence, can account for the additional uncertainty when credit cycles must be estimated from macroeconomic data.

The state-space representation of the structural UCM model that is also applied in the European system of central banks:

$$\text{Signal:} \quad y_t - y_t^* - \delta C_t = \nu_t, \nu_t \sim NID(0, \sigma^2). \quad (4.30)$$

Stochastic trend:

$$\text{State:} \quad y_t = y_{t-1}^* + \mu_{t-1} + \varepsilon_{l,t}, \varepsilon_{l,t} \sim NID(0, \sigma_l^2), \quad (4.31)$$

$$\text{State:} \quad \mu_t = m_{t-1} + \varepsilon_{\mu,t}, \varepsilon_{\mu,t} \sim NID(0, \sigma_\mu^2). \quad (4.32)$$

Stochastic cycle:

$$\text{State:} \quad C_t = c_{t-1} + \gamma_c C_{t-1}, \quad (4.33)$$

$$\text{State:} \quad (1 - 2\gamma\cos(\lambda)L + \gamma^2 L^2)c_t = (1 - \gamma\cos(\lambda)L)\varepsilon_{1,t}, \varepsilon_{1,t} \sim NID(0, 1). \quad (4.34)$$

Here:

- y_t represents the simulated time series;
- y_t^* denotes a stochastic local liner trend model consisting of level state (4.31) with variance σ_l^2 and time-varying drift μ_t (4.32) with variance σ_μ^2 ;
- ν_t is a high-frequency irregular shocks that can have varying magnitude and direction;
- c_t is the inner first-order stochastic cycle (4.34), where $\gamma \in (0, 1)$ is a damping factor of the cycle;
- $\lambda = 2\pi/T$ governs the cycle length. For example, λ of 0.2 results in a cycle length of 31 quarters corresponding to approximately 7.5-8 years;
- $Lc_t = c_{t-1}$ is a backshift operator;
- C_t is the long cycle that generates the autoregressive state equation (4.33) with dampening factor γ_c and δ in (4.30) that controls for the overall variability of the stochastic cycle.

In the simulation, all four error terms $\nu_t, \varepsilon_{l,t}, \varepsilon_{\mu,t}$ and $\varepsilon_{1,t}$ are generated as mutually and serially uncorrelated normally distributed random variables with zero mean and assessed variances. Those variances control for the signal-to-noise ratios of the stochastic cycle and the stochastic trend.

4.5 Monte Carlo method

Monte Carlo simulation is a probabilistic numerical technique widely used to estimate the outcome of an uncertain (stochastic) process. This means that it is a suitable method for simulating events that

cannot be modelled implicitly, for instance, unobserved credit cycles or replicating the distributions of ODFs similar to real credit portfolios without actual access to defaults and modelled PD data. Ease of implementation, high computing power, and cheapness are the greatest advantages of the Monte Carlo method.

In general, Monte Carlo simulations are based on the Weak Law of Large Numbers. If X_1, X_2, \dots, X_N are independent identically distributed variables with mean μ and variance σ^2 , then the term $\frac{\sigma^2}{N\varepsilon^2}$ in the Chebyshev inequality (4.35) converges to 0 when $N \rightarrow \infty$ [6].

$$\mathbb{P} \left(\left| \frac{\sum_{i=1}^N X_i}{N} - \mu \right| \geq \varepsilon \right) \leq \frac{1}{\varepsilon} E \left[\left(\frac{\sum_{i=1}^N X_i}{N} - \mu \right)^2 \right] = \frac{\sigma^2}{N\varepsilon^2}. \quad (4.35)$$

Because of the advantages mentioned above, this method will be used for our controlled simulation experiment: simulations of common component Z and portfolios.

4.6 Gaussian Mixture Model

The modelled PD distributions in real applications are often imperfectly segmented, especially when the higher granularity leads to poor quality of modelled PD. Hence, the data of the real application tend to have features (bimodality, asymmetry, shape) that are better gauged by the Gaussian mixture model. This is a probabilistic model that is based on the assumption that the data is generated using a finite number of Gaussian distributions [16]. Mixture models include latent variables indicating to which Gaussian component of the model a given point belongs.

Let latent variable be $z_i \mid \boldsymbol{\omega} \sim \text{Categorical}(\boldsymbol{\omega})$. Thus, ω_j describes the probability that z_i is equal to the j -th cluster number:

$$\mathbb{P}(z_i = j \mid \boldsymbol{\omega}) = \omega_j, \quad (4.36)$$

where $\boldsymbol{\omega} = \{\omega_1, \dots, \omega_{N_c}\}$ - vector of mixture weights, $\sum_{j=1}^{N_c} \omega_j = 1, 0 \leq \omega_j \leq 1$ and N_c is number of clusters. If $N_c = 1$ then it is simple Gaussian distribution.

The x_i are then generated from the corresponding Gaussian distribution based on the z_i values, i.e.

$$x_i \mid z_i = j \sim N(\mu_j, \sigma_j^2), \quad (4.37)$$

where μ_j, σ_j^2 - the mean and variance of the j -th cluster, $j = 1, \dots, N_c$.

In general, mixture models are based on the assumption that the data are obtained in the following way. First, the z_i variables are simulated, and then, based on the obtained z_i values, the x_i are simulated, $i = 1, \dots, N$ from the corresponding Gaussian distribution.

It follows that the density function of a Gaussian mixture model is the weighted sum of density functions of normal distributions:

$$p(x_i) = \sum_{j=1}^{N_c} \omega_j f(x_i \mid \mu_j, \sigma_j^2), \quad (4.38)$$

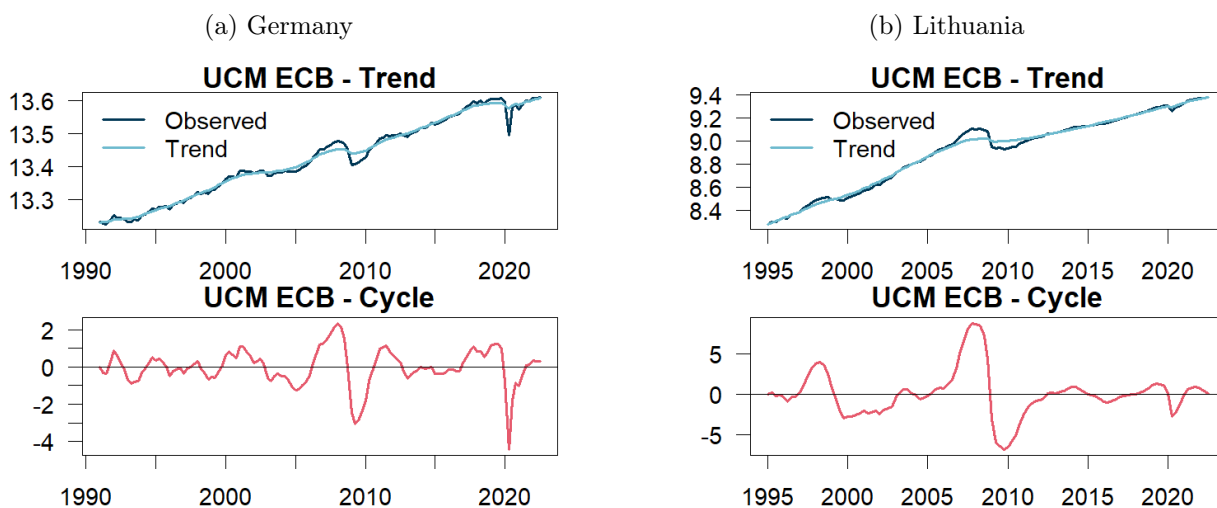
where $f(x_i | \mu_j, \sigma_j^2)$ is density function of normal distribution.

The main idea of the Gaussian mixture model is that each component of the model (cluster, class) is a separate normal distribution and each of them describes different groups of data. In the context of the thesis, clusters can be interpreted as groups of obligors making up the portfolio, e.g. better and worse customers, i.e. customers with lower and higher PDs. In our case, the Gaussian mixture will be used to generate asymmetric distribution.

4.7 Controlled simulation experiment

For the controlled simulation experiment, two countries were chosen: Lithuania and Germany. Lithuanian case represents small economies with more volatile and deeper credit cycles, while Germany represents the average of the euro area as the main contributor to the club. Common components, i.e. credit cycles, will be generated based on parameters of the UCM model obtained by applying the Bayes Markov Chain Monte Carlo (MCMC) approach to the real gross domestic product (GDP) data of these selected countries, using 22,200 simulations with 2200 used as a burn-in. Obtained parameters are provided by the supervisor and presented in Table 2. Figure 1 shows the observed time series of real GDP, extracted trend, and cycle components for Germany and Lithuania in 1a and 1b respectively.

Figure 1: Real GDP observed time series, GDP-related trend and cycle.



Posterior means from Table 2 are used to define the corresponding DE and LT Monte Carlo simulation scenarios.

4.7.1 Scenarios and assumptions

In order to examine how non-normally distributed components affect estimates of the parameters, two different scenarios are introduced:

1. Common component is generated by the stochastic simulations according to structural unobserved components model and this component is taken as given. The idiosyncratic component is

Table 2: MCMC estimated parameters of Germany and Lithuania real GDP models.

	DE real GDP, log-transformed				LT real GDP, log-transformed			
	mean	s.d.	5%	95%	mean	s.d.	5%	95%
δ	0.00852	0.00006	0.006	0.012	0.01107	0.00006	0.008	0.015
σ	0.0000486	0.0000008	0.0000150	0.000082	0.0000404	0.0000007	0.000010	0.000077
σ_l	0.0000316	0.0000008	0.0000050	0.000088	0.0000547	0.0000015	0.000007	0.000154
σ_μ	0.0000046	0.0000001	0.000002	0.000010	0.0000124	0.0000002	0.000003	0.000032
γ	0.72067	0.01008	0.461	0.919	0.82559	0.00632	0.607	0.955
γ_c	0.33405	0.01467	0.059	0.666	0.56937	0.01444	0.259	0.876
λ	0.18685	0.00548	0.064	0.378	0.19113	0.00426	0.066	0.352

generated as a normal random variable.

2. Common component is the same as in 1. The idiosyncratic component is generated as a random variable from the mixture of normal distributions.

The first scenario is considered the perfect case as Z component is assumed to be known and everything else generated as in regulations. After that, Gaussian mixture model is used for the generation of obligor-specific default threshold B_i by bringing the simulation closer to the real situation.

In the thesis, the following simplifying assumptions are made:

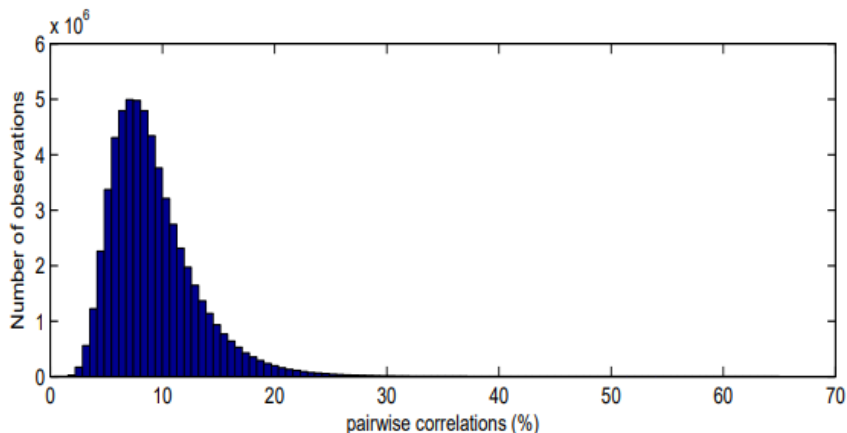
- Methodology is primarily intended to be used for corporate portfolios;
- Portfolio composition is constant;
- At every moment in the sample, an obligor can either be in default or performing independently of its previous states (multiple defaults are possible);
- Observed defaults are simulated applying PIT PD framework;
- Parameters ρ , B and α are not time-varying.

4.7.2 Algorithm of the experiment

As mentioned earlier, an experiment is implemented using Monte Carlo simulations. The subsequent algorithm aims to introduce the practicalities of the proposed estimation methodology. The first step of each simulation is to generate a common component Z for 100 time periods, i.e. 25 years of quarterly observations. For each Z , we create 5,000 realizations of idiosyncratic component ε_i . Portfolios can be of four sizes: 100, 500, 1000, and 5000 entities. To save computation time, ε_i is not regenerated each time, but the needed number of entities is randomly selected from the total set of 5,000 entities.

Then, an obligor i defaults at time t when its distance to default space $\sqrt{\rho}Z_t + \sqrt{1-\rho}\varepsilon_{i,t}$ drops below its default threshold B_i . Here, ρ is given and selected based on the distribution of pairwise correlation estimates (Figure 2) that are reported by Moody's KMV in July 2006 and related to 10,891 non-financial firms [17]. In addition, B is also given and derived from the TTC PD as in equation (4.13).

Figure 2: Distribution of pairwise correlations.



Source: Tarashev, N. and Zhu, H. (2007). Modelling and calibration errors in measures of portfolio credit risk [17].

B_i generated for each obligor and its distribution depends on the scenario. For the first scenario, B_i is generated as a Gaussian random variable $B_i \sim N(B, 1)$. In the second scenario, the distribution of B_i is constructed as a mixture of two Gaussian distributions, keeping the mean of the mixture approximately equal to B , but with the distribution being made asymmetric and skewed to the right, i.e. signifying that there are more bad clients in the portfolio (Figure 3). Parameters of the normal distributions from which the mixtures are made are selected experimentally, generating many samples according to the selected mixtures and checking their averages. Those parameters were selected with which the average of the samples generated by the mixture is closest to the true value B , but the distribution is sufficiently skewed and visually corresponds to the distributions found in reality. Parameters of Gaussian mixture components are provided in Table 3. In all cases, weights $\omega_1 = \omega_2 = 0.5$. The other parameters used for the simulation are summarized in Table 4.

Table 3: Gaussian mixture components parameters.

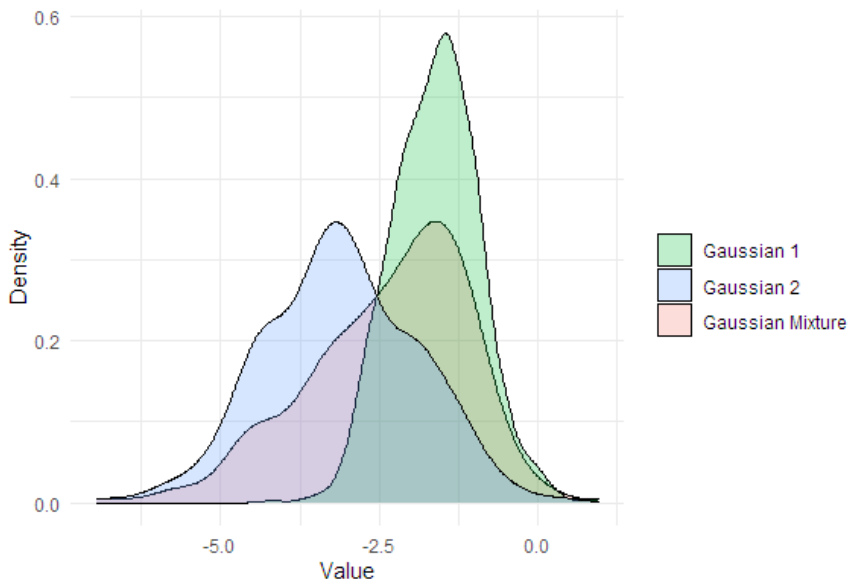
B	Gaussian components
-2.326	$x_1 \sim N(-1.59, 0.70)$, $x_2 \sim N(-3.06, 1.20)$
-2.054	$x_1 \sim N(-1.32, 0.70)$, $x_2 \sim N(-2.79, 1.20)$
-1.645	$x_1 \sim N(-0.93, 0.70)$, $x_2 \sim N(-2.36, 1.20)$
-1.282	$x_1 \sim N(-0.54, 0.70)$, $x_2 \sim N(-2.02, 1.20)$

Table 4: Simulation parameters.

Number of simulations	10,000
Time period (quarters)	100
Vasicek correlation (ρ)	0.02, 0.05, 0.10, 0.15, 0.20, 0.25
Degree of PITness (α)	0.5
TTC PD	1%, 2%, 5%, 10%
Portfolio size	100, 500, 1000, 5000

Having the number of defaulted entities in time t , the observed default frequency is calculated by dividing the number of defaults by the number of entities in the portfolio. After that, hybrid PDs for each obligor or portfolio (depending on the method) is calculated. Finally parameters ρ , B and α are estimated based on (4.21), (4.25), (4.27) and (4.29). The algorithm is repeated 10,000 times for each scenario.

Figure 3: Example of Gaussian Mixture.



5 Empirical results

5.1 Normality of simulated credit cycles

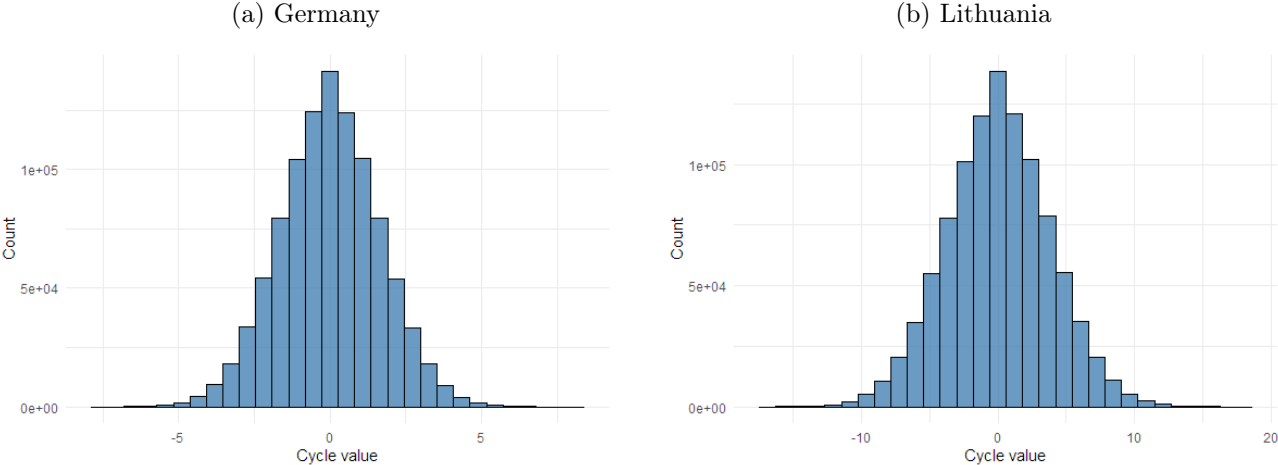
As previously mentioned, the normality hypothesis was tested for each simulation of the credit cycle (common component Z) for both simulated credit cycle cases that resemble Lithuania's and Germany's data. To do that, two well-known statistical tests were employed: the Shapiro-Wilk test and the Jarque-Bera test. The Shapiro-Wilk test is used for assessing the degree to which a dataset follows a normal distribution, based on the W statistic, which is calculated as the sum of the squares of the standardized ranks of the data. The Jarque-Bera test is another commonly used test for normality, which is based on the sample skewness and kurtosis and is often used as an alternative to the Shapiro-Wilk test in cases where the distribution of the data is unknown. In both cases, the hypotheses were tested with a significance level of $\alpha = 0.05$.

Table 5: Percentage of rejected normality hypotheses.

	Germany	Lithuania
Shapiro-Wilk test	25.2 %	50.4 %
Jarque-Bera test	11.5 %	20.2 %

The results of the tests (Table 5) suggest that the generated stochastic credit cycle is not normally distributed in a significant number of cases for both countries, Germany and Lithuania. Specifically, the Shapiro-Wilk test rejected the null hypothesis of normality in 50.4 % of the simulations for Lithuania and 25.2 % of the simulations for Germany. Similarly, the Jarque-Bera test rejected the null hypothesis of normality in 20.2 % of the simulations for Lithuania and 11.5 % of the simulations for Germany. Since the W statistic of the Shapiro-Wilk test is heavily influenced by the values of the extreme observations in the sample and is, therefore, more sensitive to deviations from normality at the tails of the distribution [4], a high percentage of hypotheses rejected by this test may (but not necessarily) indicate that there are some sharp ups and downs in the simulated cycles. For Lithuania, the Shapiro-Wilk test rejected the null hypothesis especially often, but it is quite expected because the parameters are more strict (see Table 2) and cycles are much deeper (see Figure 1b), that is, large deviations are more likely. In contrast to Shapiro-Wilk, the Jarque-Bera test is more sensitive to deviations from normality in the central part of the distribution since, as mentioned, it is based on asymmetry and peakedness of the data. Thus, the central part of the distribution of the cyclical components is more often similar to the normal distribution, i.e. Jarque-Bera test rejected the normality hypothesis less frequently for both countries.

Figure 4: Distribution of the simulated credit cycle component (combined from all simulations).



According to asymptotic theory, when the sample size becomes infinitely large, the distribution will return to normality. We decided to examine this by combining the time series from all 10,000 simulations into one long time series (1 million observations) and testing the normality of such a sample. Distributions of joined time series are presented in Figure 4. Graphical representation of the distribution of the component Z appears to be similar to the normal and it confirms the theory. However, the fact that the normality hypothesis for small samples, i.e. 100 quarterly observations were rejected for more than 5 % of all simulations means that it is rejected not randomly, i.e. non-normality is detected.

5.2 Parameter estimates

5.2.1 Scenario 1

In the methodological section, it was introduced that, in the first step, for the estimation of parameters ρ and B ODF data is needed. It may happen that, at a given time t , all obligors in the portfolio will fulfil their liabilities successfully (for example, if the portfolio consists of only good clients) and the number of observed defaults will be equal to zero. Such cases¹ in theory are called low default portfolios.

Figure 5: Percentage of simulations in which at least one point in time t the number of defaults was equal to zero: Scenario 1.

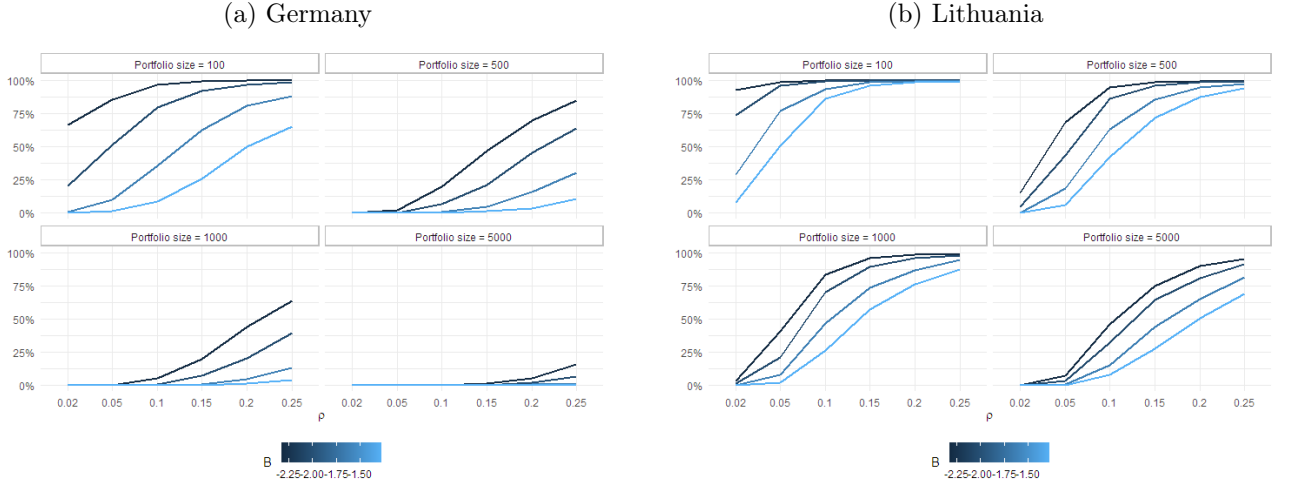


Figure 5 shows the percentage of simulations in which, for at least one time point t out of 100 observations, there were no defaults. The common tendency shows that the more negative is the value of B (i.e. lower TTC PD) and the higher is the value of ρ , the higher the percentage of cases where defaults do not occur. In other words, the portfolio of better clients, whose distance to default space is more strongly correlated with the common component Z , is more likely to be a low default portfolio. Looking at Germany and Lithuania cases separately, it could be seen that in Germany, as the size of the portfolio grows, the number of cases where there was no default decreases, and for a portfolio of size 5,000, there is a very small percentage of such cases at all. Meanwhile, in Lithuania, even for large (in the context under consideration) portfolios, the case of a low default portfolio is highly probable. It is possible that this is determined by the characteristics of the credit cycle, i.e. cycles are deeper and Z values can be higher, thus when the correlation between X_i and Z is growing it is more likely that the obligor will not default.

Unfortunately, such cases of low default portfolios complicate the evaluation of the parameters, and consequently the evaluation of the PD itself, thus additional methods should be used [5], [18]. For our chosen estimation methods, it is necessary to calculate $\Phi^{-1}(ODF_t)$, which means that the parameters ρ , B , and also α cannot be feasibly estimated when some observations are $-\infty$. Further in this thesis,

¹Also when the number of defaults is not zero, but only a few default observations occur.

the averages of the parameter estimates and their interpretations will be presented for all the examined cases², but it is important to keep in mind that, in some cases, the parameters were not estimated for a large number of simulations and the average is calculated for a relatively small sample, so the estimates may be significantly biased and unreliable.

Continuing on defaults, Table 9 shows that the percentage of defaults in the portfolio does not depend on the size of the portfolio. The ODF varies only through different values of ρ and B . For Germany, the ODF varies little across different ρ values. The difference between the ODF values when ρ takes its maximum and minimum values (i.e. the difference between the rightmost and leftmost values of Table 9) varies between 1.6 and 2.1 percentage points. Meanwhile, for Lithuania, this difference ranges between 9.0 - 9.8 p.p. Additionally, when comparing the countries, in the case of Lithuania, the ODF is higher than that of Germany in all cells. It is interesting that Lithuania has more cases of low default portfolios, but when they are excluded, the ODFs for Lithuania are larger.

Figure 6: Deviations of the parameter ρ estimates from the true values of the parameter: Scenario 1.

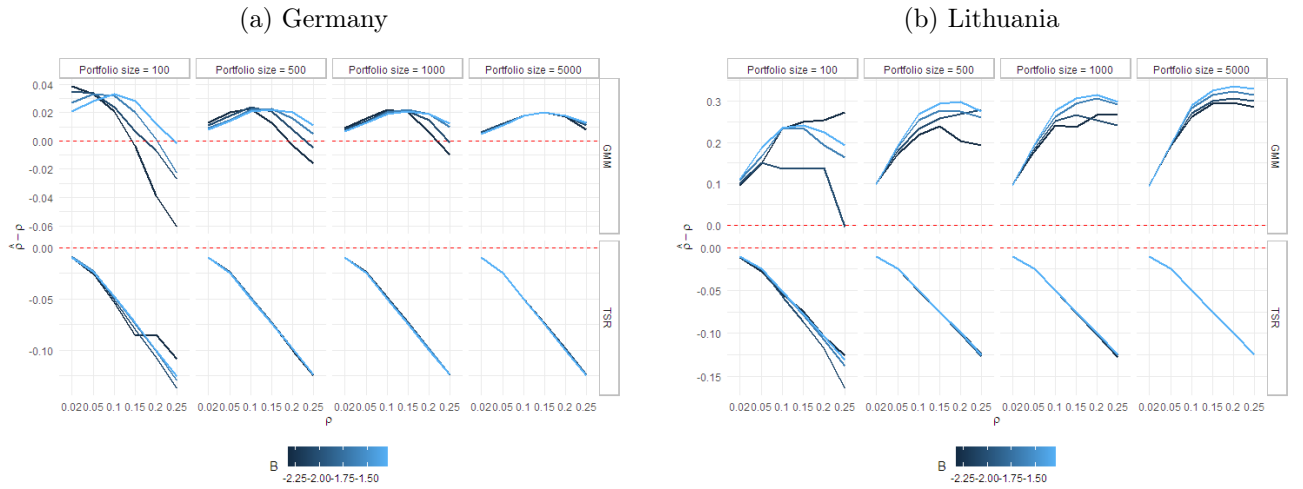


Figure 6 presents deviations of the estimated parameter ρ from its true value, i.e. the difference between the estimate and true value of the parameter, while Table 6 presents estimates of ρ . In the case of Germany, when the true value of ρ is small (e.g. 2 % or 5 %), the two-step regression estimates ρ fairly accurately, but as the value of ρ increases, the deviation from the true value increases rapidly. GMM estimates the ρ parameter much more accurately and stably for Germany. Figure 6b shows quite the opposite situation in the case of Lithuania. Here GMM loses its stability and differences start to increase for higher ρ values. Thus, the TSR method works quite similarly in both cases while GMM performs poorly in the case of Lithuania, which means that the not normally distributed common component Z affects estimates obtained by the GMM method. Additionally, when the number of obligors in the portfolio is growing, the true value of B no longer affects $\hat{\rho}$, i.e. lines in the graphs overlap. Although this also applies to the GMM method in the case of Lithuania, it can be seen (top row of Figure 6b) that for a larger B , i.e. for the portfolio of worse clients, ρ is more overestimated in

²Except two cases of Lithuania in Scenario 2., when $\rho = 0.25$, $\text{TTC PD} \in \{1\%, 2\%\}$ and portfolio size = 100, because 100% of simulations were zero default cases.

the GMM case.

Parameter B is overestimated by both methods and for both countries, i.e. we consider the portfolio worse than it actually is (Figure 7 and Table 7). As the true value of B increases, the deviations of the estimates from it decrease slightly, i.e. for poorer portfolios, we can more accurately estimate the parameter. In the case of two-step regression, all the blue lines in Figure 7 are quite close to each other, which means that the accuracy of \hat{B} does not vary strongly for different values of ρ . Whereas, in the GMM method, the estimate of B for different values of ρ is quite different, especially for Lithuania.

The last parameter assessed in Scenario 1. is α . Estimates of this parameter are presented in Table 8 and visually plotted in Figure 8. The true value of α is equal to 0.05 and is marked in the graphs as a red dashed line. It can be seen, that TSR overestimates it for both Germany and Lithuania, and true values of ρ and B do not affect the $\hat{\alpha}$ in this method. GMM performs quite accurately for Germany but underestimates the parameter for Lithuania. Since for the assessment of α an estimate of ρ is used and, as discussed, is overestimated, this potentially affects the underestimation of $\hat{\alpha}$.

5.2.2 Scenario 2

In the second scenario, everything remains as in the first, but the distribution of B_i is changed by creating asymmetry and skewing the distribution to the right using the Gaussian mixture method. The obtained results are presented analogously as before.

Percentages of no default portfolio cases are presented in Figure 15. The graphs show that the results are very similar to the results of the first scenario. Since B_i is shifted to the right, it was expected that there would be fewer such portfolios where no obligor defaults. It also follows that ODFs in Scenario 2. increase by approximately 1 - 3 p.p. (see Table 16).

When estimating the parameter ρ , the overall trends remain the same: GMM method overestimates and TSR underestimates $\hat{\rho}$, but the main difference obtained between scenarios is the dependence of the accuracy of parameter estimates on the true value of parameter B (see Figure 12). The difference in GMM is more clearly visible, and TSR is less but compared to Scenario 1., there is a change. In the case of Germany, dependence on B became evident, i.e. the wider spread between lines, while the lines in Lithuania's graphs overlap more, so when we make B_i as a Gaussian mixture, the influence of the actual value of B decreases. Interestingly, the GMM for Germany estimates ρ more accurately when the TTC PD is higher (light blue lines in Figure 12a are lower). The opposite situation occurs for Lithuania and the deviations are smaller, when the TTC PD is lower (see Figure 12b).

Figure 13 shows that the generation of B_i as a mixture did not have a significant effect on the estimation of the B parameter by the GMM method. The only slightly more pronounced change is that when TTC PD is higher, GMM assesses more similarly for different ρ . TSR with B_i as a mixture performs worse in the evaluation, deviations increased both in Germany's and Lithuania's cases. In the case of Germany, GMM and TSR perform very similarly, while in Lithuania's case, there are greater differences between these methods.

Finally, estimates of the parameter α for Scenario 2. are presented in Figure 14 and Table 15. Consistent with the already discussed results of Scenario 1., the GMM method underestimates and

TSR overestimates the value of $\hat{\alpha}$. As in the case of other parameter estimates, larger differences appear between various variants of B (except for Lithuania’s GMM case).

5.2.3 Hybrid estimation method

The results of the first and second scenarios revealed that, in general, the GMM method overestimates the ρ parameter, while the TSR underestimates it. For the α parameter, the methods work the other way around: GMM underestimates, while TSR overestimates. This suggests that the methods seem to work in opposite directions, and from this follows the possible idea that the combination of GMM and TSR would be a better hybrid approach for TTC parameters estimation. To test this hypothesis, Monte Carlo simulations were repeated for the previously generated Z values, slightly changing the principle of parameter estimation. Initially, $\hat{\rho}_{GMM}, \hat{\rho}_{TSR}, \hat{B}_{GMM}$ and \hat{B}_{TSR} were estimated by GMM and TSR methods as before. New estimates of these parameters then were calculated as the average of the estimates obtained by the different methods:

$$\hat{\rho}_{hybrid} = \frac{\hat{\rho}_{GMM} + \hat{\rho}_{TSR}}{2} \quad \text{and} \quad \hat{B}_{hybrid} = \frac{\hat{B}_{GMM} + \hat{B}_{TSR}}{2}. \quad (5.1)$$

After that, the α parameter was again estimated by two methods, GMM and TSR according to equations (4.25) and (4.29), but using the hybrid ρ estimate $\hat{\rho}_{hybrid}$. Finally, a hybrid estimate of α was also obtained by averaging $\hat{\alpha}_{GMM}$ and $\hat{\alpha}_{TSR}$:

$$\hat{\alpha}_{hybrid} = \frac{\hat{\alpha}_{GMM} + \hat{\alpha}_{TSR}}{2}. \quad (5.2)$$

The estimates of the parameters obtained in Scenario 1. are given in Tables 10, 11, 12 and graphically represented in Figures 9, 10 and 11 for ρ , B and α respectively. Since for Germany, the estimates of the ρ and α parameters for different methods had deviations in different directions at a similar distance for small ρ values, the hybrid method, when the estimates are averaged across methods, works well and gives relatively accurate estimates of ρ and α (see Figure 9a and Figure 11a). However, for larger true values of ρ , severely underestimated TSR estimates weigh the average down, and deviations from the true values of the hybrid estimates remain large. The accuracy of estimates of the parameter B was not improved much by the hybrid method. In the case of Lithuania, GMM overestimates ρ more strongly than TSR underestimates it, therefore, when averaging, $\hat{\rho}_{hybrid}$ remains significantly overestimated and, consequently, $\hat{\alpha}_{hybrid}$ - underestimated, as is shown in Figure 9b and Figure 11b.

The results of Scenario 2. are presented analogously in tables (Tables 17, 18, 19) and graphs (Figures 16, 17, 18). The insights that follow are broadly similar to those of Scenario 1., but here the case of Germany stands out with relatively accurate estimates of the parameters ρ and α . As for the GMM and TSR methods in this scenario, there also remains a variation in estimates for different actual values of B due to the simulation of B_i as a Gaussian mixture.

6 Conclusions

The normality hypothesis for small samples with 100 quarterly observations was rejected in more than 5% of all simulations for both countries, Germany and Lithuania, indicating that non-normality is present in data samples similar to real applications and is not just occurring randomly. This indicates, that if the credit cycle is used as a common component in the ASRF model, the assumption that the systematic factor is distributed according to the standard normal distribution is likely to be violated. For the Lithuania case, the percentage of the rejected hypothesis was especially high and as further results have shown, this has a rather significant impact on the accuracy of parameter estimation.

The obtained parameter estimation results of Scenario 1. showed that the deviations of the common component from normality significantly affect the GMM results. The estimates of both main TTC PD parameters, Vasicek's correlation and degree of PITness, were significantly more accurate for Germany than for Lithuania when evaluated by the GMM method. Meanwhile, the TSR, in the cases of both Lithuania and Germany, evaluated underlying parameters much more inaccurately than the GMM, but almost identically wrong in the cases of both countries. That means that the non-normality of the common component has not detrimentally affected the TSR estimates.

Since the data of the real application tend to have features, such as asymmetry or shape, in Scenario 2., the obligor-specific default threshold, which accounts for idiosyncratic effects, was generated according to the Gaussian mixture model. The main difference between the results of the scenarios is the dependence of the accuracy of parameter estimates on the quality of the portfolio.

The comparison of the two methods showed that the GMM method tends to overestimate, while the TSR tends to underestimate when evaluating the Vasicek correlation. When estimating the PITness parameter, the methods work in the opposite way - GMM underestimates and TSR overestimates the parameter. However, this is directly related to the deviation of the correlation parameter estimate from its actual value, since both methods use the estimate of the first parameter to assess the PITness parameter non-normality of the common component has not affected the estimates.

Having noticed that the GMM and TSR methods work in "different directions", it was decided to test a combination of these two by introducing a hybrid method that takes an average of GMM and TSR parameter estimates. The results showed that such a method does not work well for the case of Lithuania neither in Scenario 1. nor Scenario 2. because GMM overestimates the ρ parameter too much, thus the hybrid estimate remains overestimated. In the case of Germany, the hybrid method performed well for lower actual values of Vasicek correlation but the accuracy dramatically decreased for higher ones. However, in Scenario 2., the combination of GMM and TSR provides relatively accurate estimates of the Vasicek correlation and PITness parameter even for the higher true values of ρ . To sum up, such a method could also be possibly used for the estimation of TTC PD parameters for comparison with other methods and the actual averaging scheme could be farther elaborated.

Finally, the results highlighted the importance of the low default portfolio problem that is quite often encountered in reality. GMM and TSR methods employ ODF in the parameter estimation, thus, when the number of defaults in the portfolio is equal to zero for at least one point time, the TTC PD parameters cannot be feasibly estimated. Results have shown that the portfolio of better clients, as

well as those whose distance to default space is more strongly correlated with the common component, is more likely to observe zero defaults. This means that it is important to know the characteristics of the portfolio before choosing a valuation method. The problem is also well-understood by regulators, as the forthcoming Basel IV will no longer allow low default portfolios to use internal rating models.

In the context of this thesis, a common component, i.e. credit cycle, is considered as known, but in real-life applications, the credit cycle is not observed, so it must be extracted from macroeconomic data, e.g. by decomposing it or other methods. Since our data is simulated we have full knowledge of simulated elements, i.e. cycle, trend, and noise, and can account for the additional uncertainty. Therefore, for the extension of the thesis, it would be possible to extract the cycle component from the simulated macroeconomic data using e.g. Hodrick-Prescott or other filters, compare it to the actual cyclical component and assess how much the cycle estimation errors, in such case, would contribute to TTC PD parameters estimates.

References

- [1] S. D. Aguais, Forest Jr. L. R., M. King, M. C. Lennon, and B. Lordkipanidze. Designing and implementing a basel ii compliant pit–ttc ratings framework. *The Basel Handbook: A Guide for Financial Practitioners*, pages 267–297, 2008.
- [2] M. Carlehed and A. Petrov. A methodology for point-in-time–through-the-cycle probability of default decomposition in risk classification systems. *The Journal of Risk Model Validation*, 6(3):3–25, 2012.
- [3] D. Celov and M. Comunale. Business cycles in the eu: A comprehensive comparison across methods. *Bank of Lithuania. Discussion paper series 26*, 2021.
- [4] R. Cont, editor. *Statistical Tests for Normality*. John Wiley & Sons, Inc, 2010.
- [5] L. Gerhardsson and Castor N. Estimation of probability of default in low default portfolios. 2017.
- [6] M. Gilli, D. Maringer, and E. Schumann. Generating random numbers. *Numerical Methods and Optimization in Finance*, pages 103–132, 2019.
- [7] J. Happonen. Theory behind regulatory capital formulae. *Master’s Thesis, Aalto University*, 2016.
- [8] B. Oeyen and O. S. Celis. On probability of default and its relation to observed default frequency and a common factor. *Journal of Credit Risk*, 15(3):41–66, 2019.
- [9] Basel Committee on Banking Supervision. International convergence of capital measurement and capital standards. *Bank for International Settlements*, 1988.
- [10] Basel Committee on Banking Supervision. International convergence of capital measurement and capital standards: A revised framework. *Bank for International Settlements*, 2004.
- [11] Basel Committee on Banking Supervision. International convergence of capital measurement and capital standards: A revised framework. *Bank for International Settlements*, 2006.
- [12] Basel Committee on Banking Supervision. Basel iii: A global regulatory framework for more resilient banks and banking systems. *Bank for International Settlements*, 2010.
- [13] Basel Committee on Banking Supervision. Basel iii: The liquidity coverage ratio and liquidity risk monitoring tools. *Bank for International Settlements*, 2013.
- [14] Basel Committee on Banking Supervision. Basel iii: The net stable funding ratio. *Bank for International Settlements*, 2016.
- [15] M. Rubtsov and A. Petrov. A point-in-time–through-the-cycle approach to rating assignment and probability of default calibration. *The Journal of Risk Model Validation*, 10(2):83–112, 2016.

- [16] C. Rudin. Gaussian mixture models and expectation maximization. <https://users.cs.duke.edu/~cynthia/teaching.html>, 2021.
- [17] N. Tarashev and H. Zhu. Modelling and calibration errors in measures of portfolio credit risk. *Bank for International Settlements. Working Papers*, No. 230, 2007.
- [18] D. Tasche and K. Pluto. Estimating probabilities of default for low default portfolios. *Banking & Insurance*, 2004.
- [19] O. A. Vasicek. The distribution of loan portfolio value. *Risk*, 15(12):160–162, 2002.

7 Appendix A. Figures and tables

Figure 7: Deviations of the parameter B estimates from the true values of the parameter: Scenario 1.

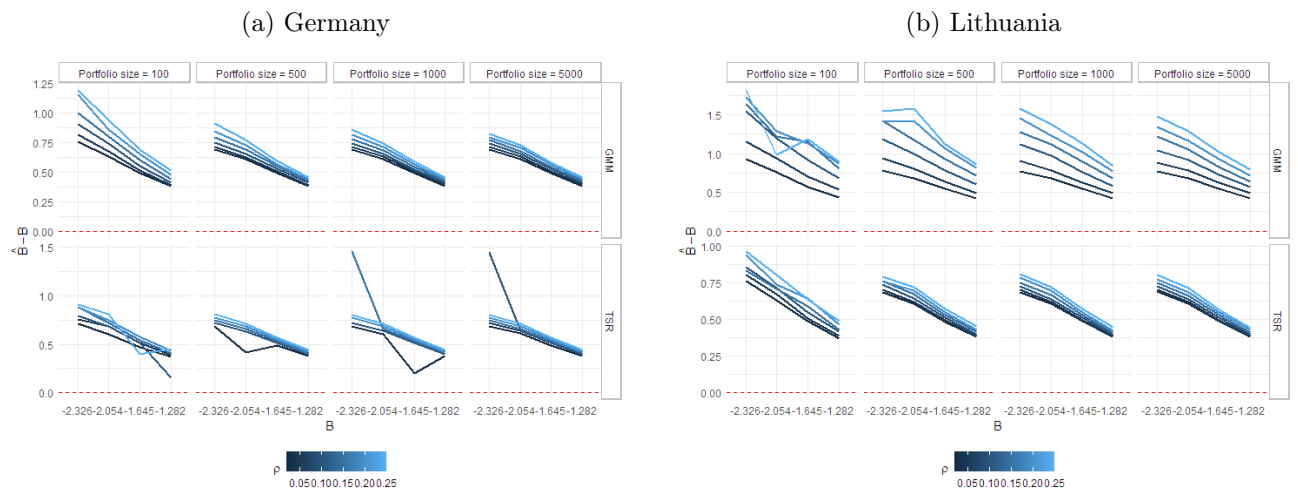


Figure 8: Estimates of the parameter α : Scenario 1.

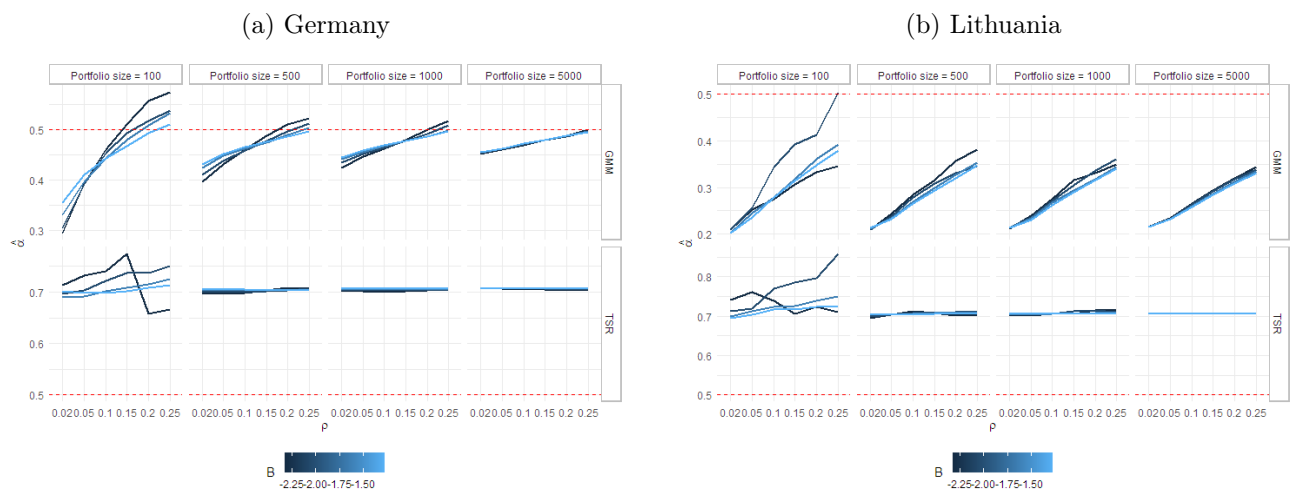


Figure 9: Deviations of the parameter ρ hybrid estimates from the true values of the parameter: Scenario 1.

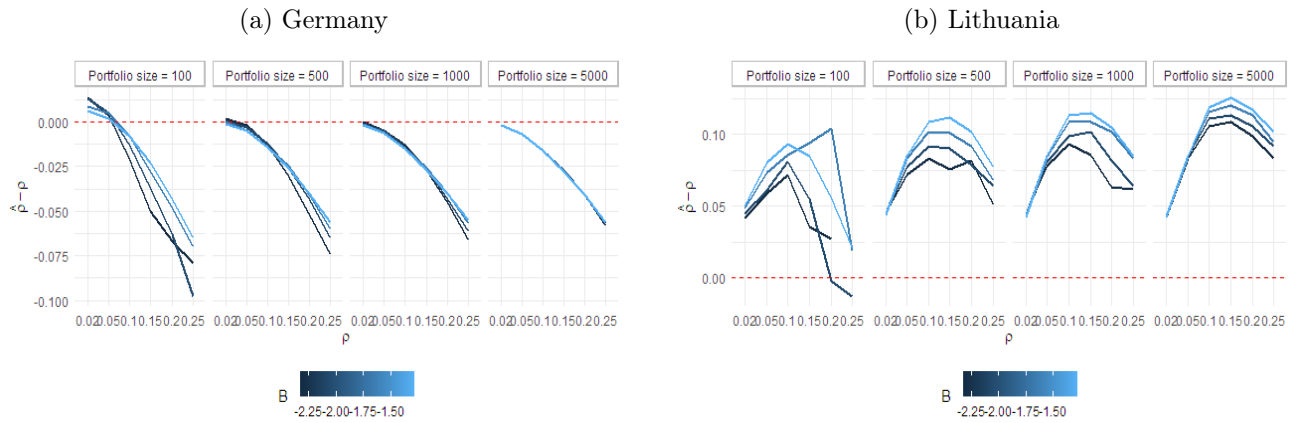


Figure 10: Deviations of the parameter B hybrid estimates from the true values of the parameter: Scenario 1.

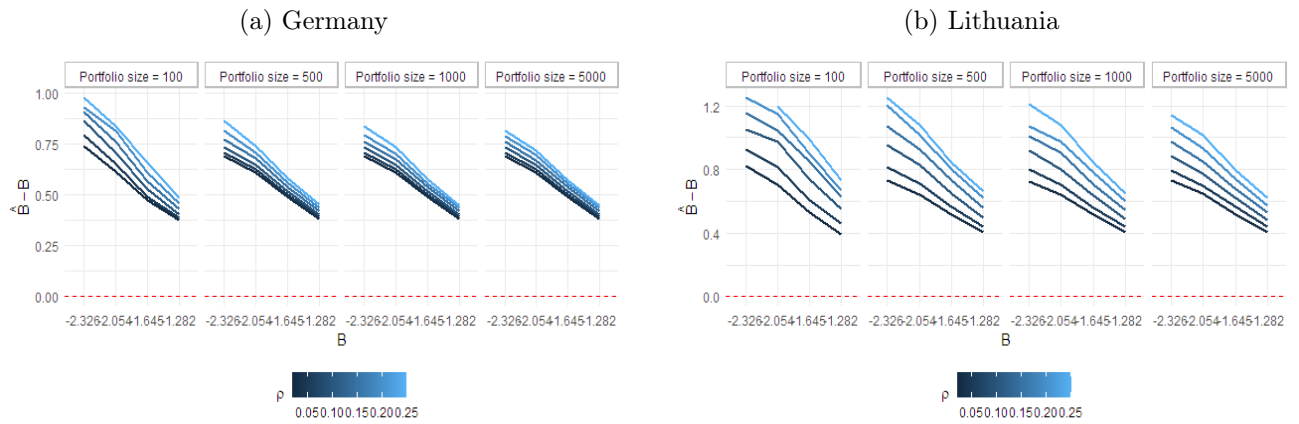


Figure 11: Deviations of the parameter α hybrid estimates from the true values of the parameter: Scenario 1.

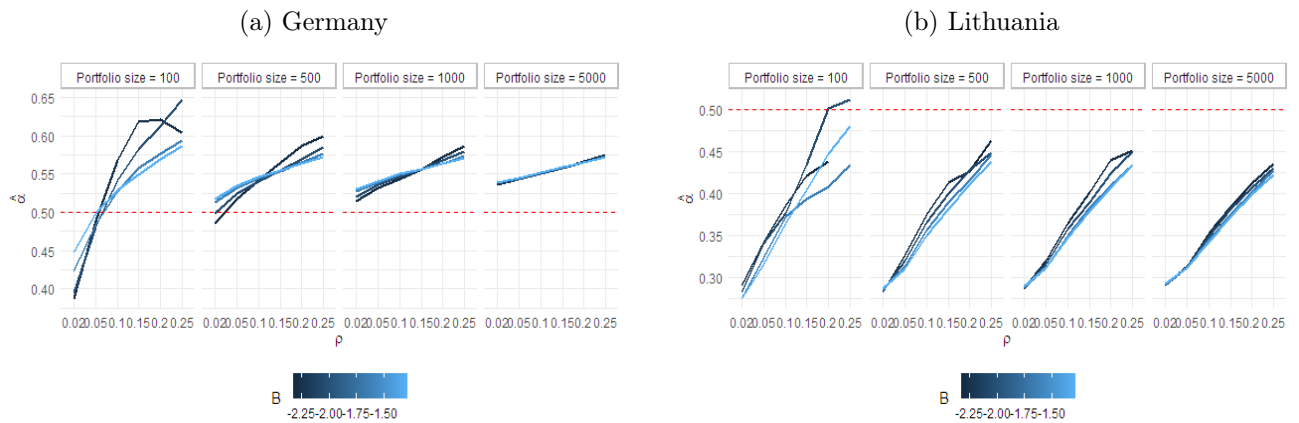


Figure 12: Deviations of the parameter ρ estimates from the true values of the parameter: Scenario 2.

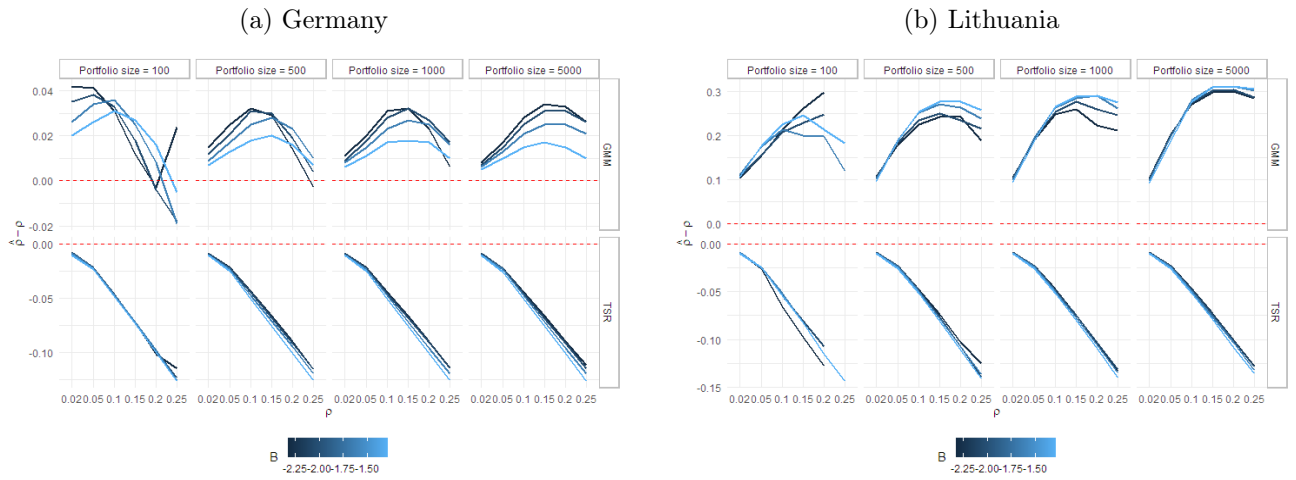


Figure 13: Deviations of the parameter B estimates from the true values of the parameter: Scenario 2.

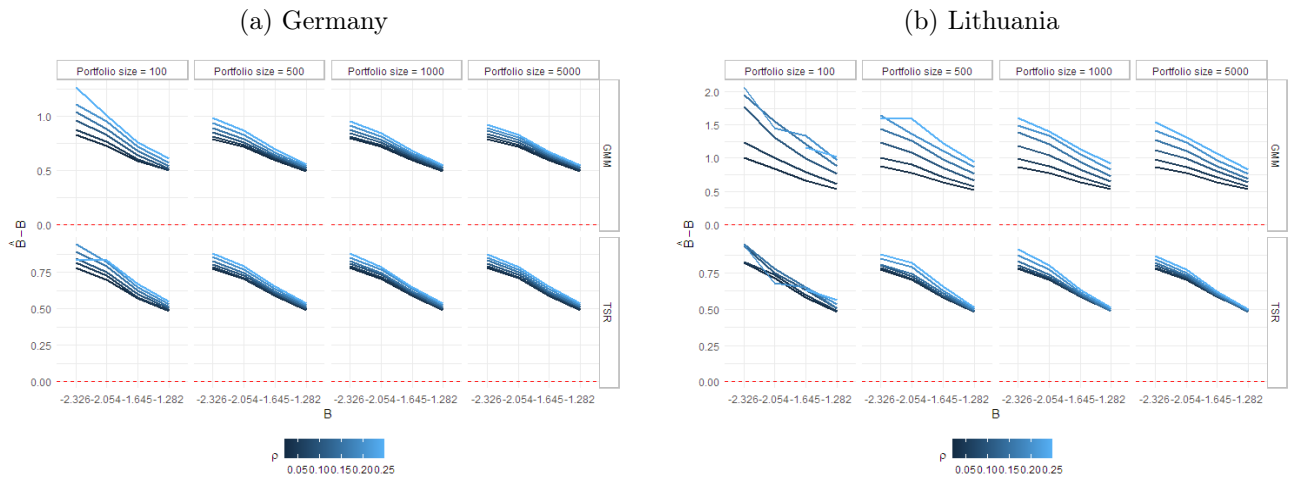


Figure 14: Estimates of the parameter α : Scenario 2.

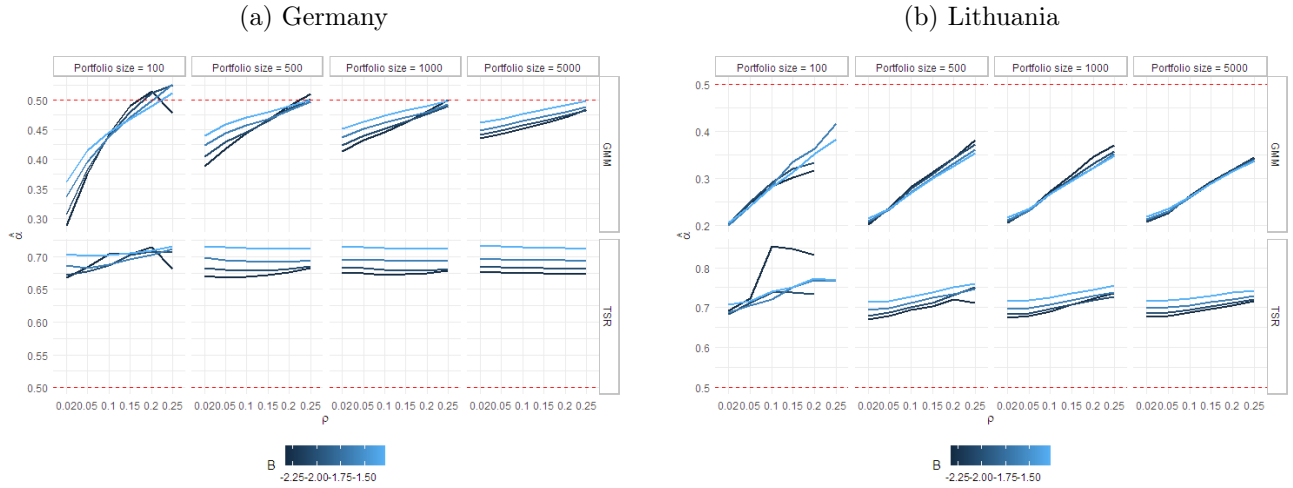


Figure 15: Percentage of simulations in which at at least one point in time t the number of defaults was equal to zero: Scenario 2.

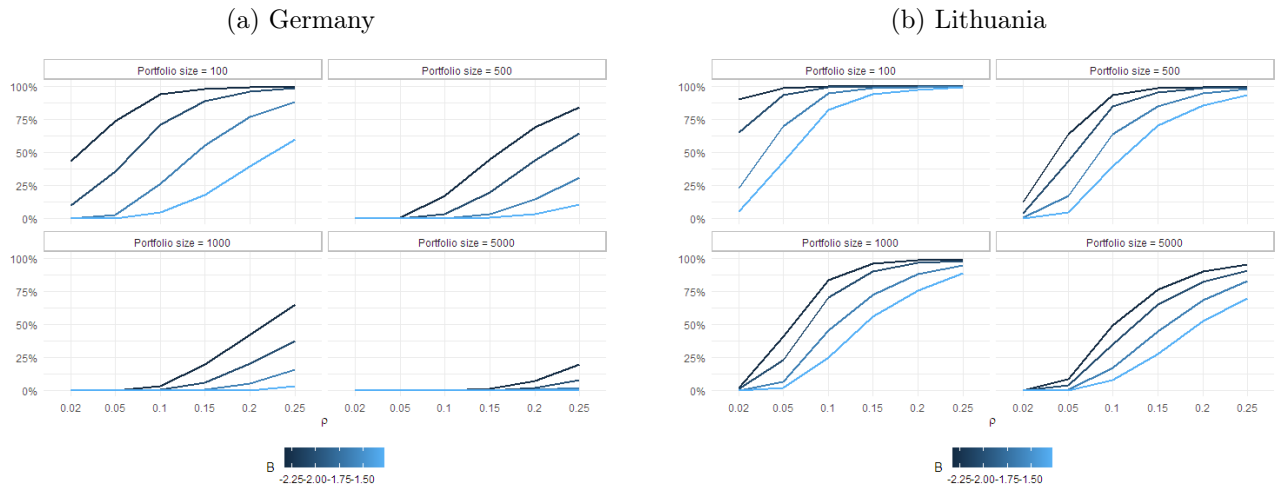


Figure 16: Deviations of the parameter ρ hybrid estimates from the true values of the parameter: Scenario 2.

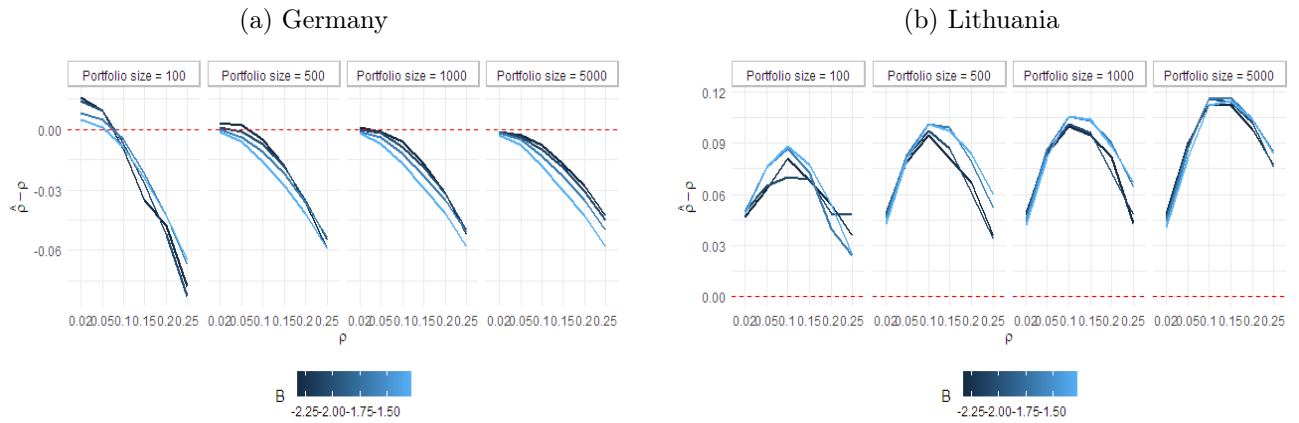


Figure 17: Deviations of the parameter B hybrid estimates from the true values of the parameter: Scenario 2.

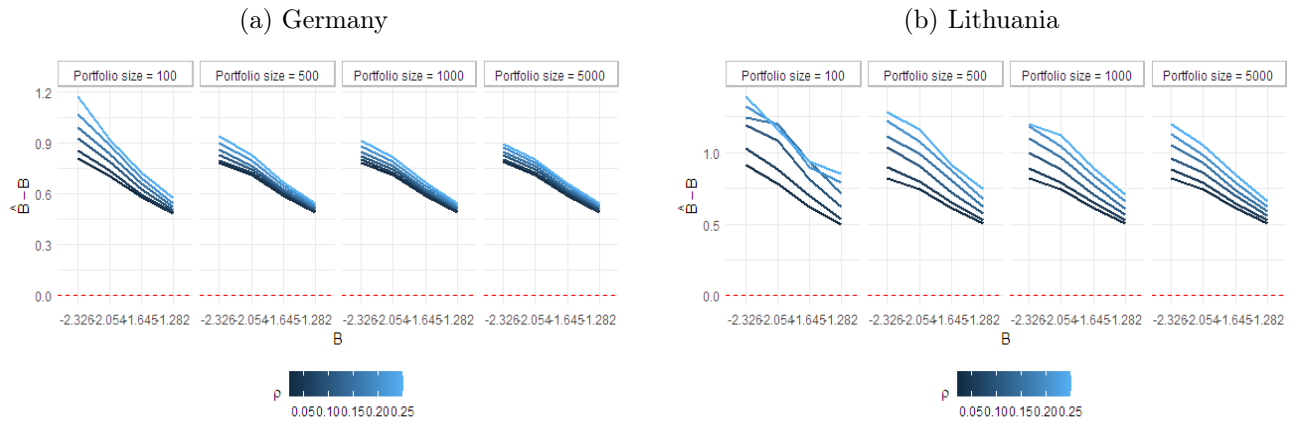


Figure 18: Deviations of the parameter α hybrid estimates from the true values of the parameter: Scenario 2.

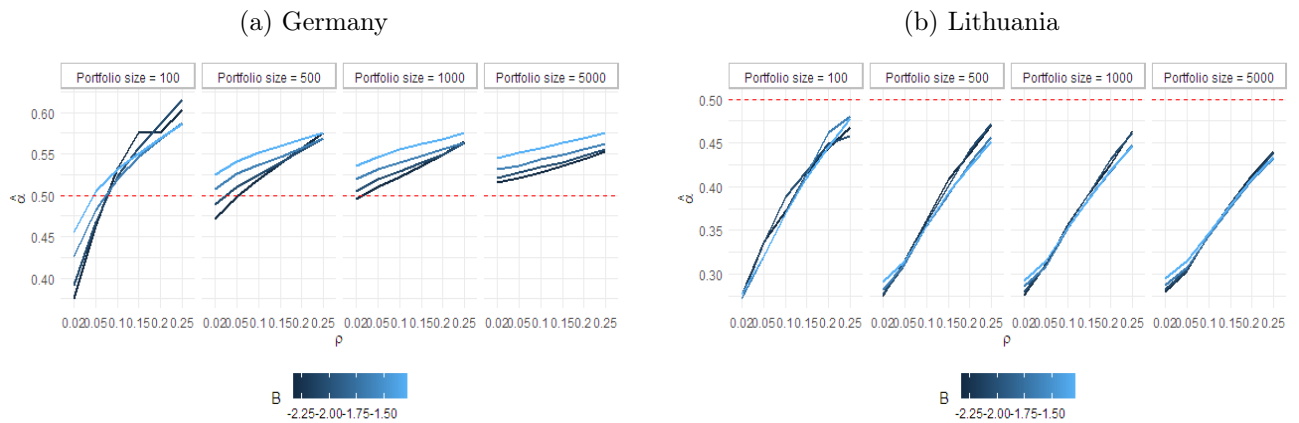


Table 6: Estimates of Vasicek correlation parameter ρ : Scenario 1.

Portfolio size	ρ	0.02	0.05	0.10	0.15	0.20	0.25
	B	GMM TSR	GMM TSR	GMM TSR	GMM TSR	GMM TSR	GMM TSR
Germany							
100	-2.326	0.059 0.010 (0.666)	0.083 0.024 (0.852)	0.121 0.046 (0.963)	0.146 0.064 (0.992)	0.161 0.115 (0.999)	0.189 0.141 (0.999)
	-2.054	0.055 0.011 (0.204)	0.084 0.026 (0.511)	0.124 0.049 (0.792)	0.157 0.070 (0.917)	0.193 0.093 (0.969)	0.223 0.112 (0.986)
	-1.645	0.047 0.011 (0.007)	0.083 0.027 (0.098)	0.132 0.051 (0.352)	0.170 0.075 (0.627)	0.201 0.099 (0.805)	0.227 0.120 (0.879)
	-1.282	0.041 0.010 (0.000)	0.078 0.026 (0.014)	0.133 0.052 (0.086)	0.178 0.077 (0.256)	0.212 0.100 (0.500)	0.248 0.124 (0.651)
500	-2.326	0.033 0.010 (0.001)	0.070 0.026 (0.020)	0.123 0.051 (0.198)	0.163 0.076 (0.465)	0.197 0.100 (0.694)	0.234 0.125 (0.849)
	-2.054	0.031 0.010 (0.000)	0.068 0.026 (0.002)	0.124 0.051 (0.062)	0.171 0.077 (0.210)	0.208 0.101 (0.452)	0.245 0.126 (0.637)
	-1.645	0.029 0.010 (0.000)	0.065 0.025 (0.000)	0.122 0.051 (0.008)	0.173 0.076 (0.047)	0.216 0.101 (0.155)	0.255 0.126 (0.300)
	-1.282	0.028 0.010 (0.000)	0.064 0.025 (0.000)	0.121 0.050 (0.000)	0.172 0.076 (0.010)	0.220 0.101 (0.033)	0.261 0.126 (0.107)
1000	-2.326	0.029 0.010 (0.000)	0.066 0.026 (0.002)	0.122 0.051 (0.052)	0.170 0.077 (0.196)	0.206 0.101 (0.441)	0.240 0.125 (0.640)
	-2.054	0.028 0.010 (0.000)	0.064 0.025 (0.000)	0.121 0.051 (0.003)	0.171 0.076 (0.073)	0.215 0.101 (0.204)	0.249 0.126 (0.395)
	-1.645	0.027 0.010 (0.000)	0.063 0.025 (0.000)	0.120 0.050 (0.002)	0.172 0.076 (0.008)	0.219 0.101 (0.043)	0.260 0.126 (0.132)
	-1.282	0.027 0.010 (0.000)	0.063 0.025 (0.000)	0.119 0.050 (0.000)	0.171 0.075 (0.002)	0.219 0.100 (0.009)	0.262 0.126 (0.041)
5000	-2.326	0.026 0.010 (0.000)	0.062 0.025 (0.000)	0.118 0.050 (0.002)	0.170 0.076 (0.014)	0.217 0.101 (0.052)	0.258 0.126 (0.157)
	-2.054	0.026 0.010 (0.000)	0.062 0.025 (0.000)	0.118 0.050 (0.000)	0.170 0.075 (0.001)	0.218 0.101 (0.016)	0.261 0.126 (0.062)
	-1.645	0.025 0.010 (0.000)	0.061 0.025 (0.000)	0.118 0.050 (0.000)	0.170 0.075 (0.000)	0.218 0.100 (0.005)	0.262 0.125 (0.015)
	-1.282	0.025 0.010 (0.000)	0.061 0.025 (0.000)	0.118 0.050 (0.000)	0.170 0.075 (0.000)	0.218 0.100 (0.001)	0.263 0.125 (0.003)
Lithuania							
100	-2.326	0.117 0.009 (0.926)	0.199 0.022 (0.988)	0.333 0.046 (0.996)	0.400 0.076 (0.998)	0.454 0.096 (0.999)	0.523 0.124 (0.999)
Continued on next page							

Table 6 – continued from previous page

Portfolio size	ρ	0.02		0.05		0.10		0.15		0.20		0.25	
	B	GMM TSR		GMM TSR		GMM TSR		GMM TSR		GMM TSR		GMM TSR	
100	-2.054	0.121	0.010	0.200	0.024	0.236	0.043	0.286	0.063	0.338	0.082	0.245	0.085
		(0.733)		(0.960)		(0.995)		(0.998)		(0.998)		(0.999)	
	-1.645	0.128	0.010	0.216	0.025	0.334	0.048	0.385	0.072	0.393	0.092	0.413	0.112
		(0.287)		(0.769)		(0.936)		(0.986)		(0.995)		(0.997)	
	-1.282	0.130	0.010	0.237	0.025	0.336	0.049	0.392	0.073	0.424	0.096	0.444	0.119
		(0.079)		(0.506)		(0.858)		(0.957)		(0.985)		(0.993)	
500	-2.326	0.121	0.010	0.221	0.025	0.318	0.049	0.390	0.075	0.404	0.101	0.443	0.126
		(0.151)		(0.681)		(0.946)		(0.987)		(0.994)		(0.996)	
	-2.054	0.121	0.010	0.231	0.025	0.333	0.050	0.408	0.075	0.469	0.099	0.529	0.123
		(0.043)		(0.434)		(0.860)		(0.957)		(0.986)		(0.992)	
	-1.645	0.120	0.010	0.239	0.025	0.355	0.050	0.427	0.075	0.476	0.100	0.510	0.124
		(0.002)		(0.186)		(0.632)		(0.857)		(0.945)		(0.971)	
	-1.282	0.119	0.010	0.244	0.025	0.368	0.050	0.444	0.075	0.499	0.101	0.526	0.125
		(0.001)		(0.061)		(0.420)		(0.715)		(0.874)		(0.938)	
1000	-2.326	0.119	0.010	0.229	0.025	0.341	0.050	0.388	0.075	0.467	0.098	0.518	0.122
		(0.031)		(0.405)		(0.836)		(0.958)		(0.984)		(0.990)	
	-2.054	0.118	0.010	0.235	0.025	0.352	0.050	0.416	0.074	0.454	0.098	0.491	0.124
		(0.009)		(0.209)		(0.700)		(0.895)		(0.959)		(0.982)	
	-1.645	0.118	0.010	0.243	0.025	0.363	0.050	0.445	0.075	0.506	0.100	0.542	0.125
		(0.001)		(0.075)		(0.464)		(0.733)		(0.868)		(0.948)	
	-1.282	0.117	0.010	0.246	0.025	0.378	0.050	0.456	0.075	0.514	0.100	0.548	0.125
		(0.000)		(0.016)		(0.259)		(0.569)		(0.765)		(0.873)	
5000	-2.326	0.116	0.010	0.241	0.025	0.363	0.050	0.444	0.075	0.496	0.100	0.536	0.125
		(0.000)		(0.068)		(0.460)		(0.750)		(0.898)		(0.955)	
	-2.054	0.116	0.010	0.244	0.025	0.371	0.050	0.451	0.075	0.507	0.100	0.551	0.125
		(0.000)		(0.032)		(0.320)		(0.642)		(0.808)		(0.910)	
	-1.645	0.116	0.010	0.245	0.025	0.383	0.050	0.465	0.075	0.523	0.100	0.566	0.125
		(0.000)		(0.008)		(0.151)		(0.439)		(0.651)		(0.813)	
	-1.282	0.115	0.010	0.245	0.025	0.389	0.050	0.476	0.075	0.537	0.100	0.580	0.125
		(0.000)		(0.002)		(0.075)		(0.272)		(0.504)		(0.690)	

Note: in the parentheses below estimates is specified the ratio of simulations in which at at least one point in time t the number of defaults was equal to zero.

Table 7: Estimates of parameter B : Scenario 1.

Portfolio size	ρ	0.02		0.05		0.10		0.15		0.20		0.25	
	B	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR
Germany													
100	-2.326	-1.568	-1.612	-1.509	-1.569	-1.421	-1.531	-1.324	-1.443	-1.168	-1.445	-1.128	-1.415
	-2.054	-1.424	-1.454	-1.378	-1.374	-1.311	-1.374	-1.251	-1.335	-1.192	-1.306	-1.113	-1.248
	-1.645	-1.155	-1.173	-1.133	-1.131	-1.090	-1.131	-1.044	-1.101	-0.989	-1.067	-0.956	-1.248
	-1.282	-0.903	-0.914	-0.891	-0.891	-0.869	-1.131	-0.842	-0.874	-0.806	-0.852	-0.771	-0.831
500	-2.326	-1.632	-1.643	-1.611	-1.607	-1.573	-1.607	-1.529	-1.578	-1.481	-1.550	-1.413	-1.520
	-2.054	-1.443	-1.643	-1.426	-1.421	-1.396	-1.421	-1.362	-1.397	-1.328	-1.372	-1.281	-1.345
	-1.645	-1.153	-1.159	-1.140	-1.134	-1.119	-1.134	-1.097	-1.117	-1.075	-1.099	-1.046	-1.078
	-1.282	-0.900	-0.904	-0.890	-0.885	-0.873	-0.885	-0.858	-0.872	-0.843	-0.858	-0.824	-0.842
1000	-2.326	-1.632	-1.640	-1.613	-1.606	-1.582	-1.606	-1.547	-0.872	-1.506	-1.552	-1.462	-1.522
	-2.054	-1.441	-1.447	-1.424	-1.416	-1.398	-1.416	-1.371	-1.395	-1.341	-1.372	-1.310	-1.347
	-1.645	-1.154	-1.447	-1.141	-1.134	-1.120	-1.134	-1.100	-1.117	-1.080	-1.098	-1.056	-1.079
	-1.282	-0.899	-0.903	-0.889	-0.883	-0.872	-0.883	-0.857	-0.870	-0.842	-0.856	-0.826	-0.840
5000	-2.326	-1.632	-1.638	-1.613	-1.602	-1.584	-0.883	-1.556	-1.578	-1.529	-1.552	-1.499	-1.525
	-2.054	-1.440	-1.445	-1.424	-1.414	-1.398	-1.414	-1.374	-1.393	-1.350	-1.370	-1.326	-1.345
	-1.645	-1.153	-1.157	-1.140	-1.132	-1.119	-1.132	-1.100	-1.115	-1.082	-1.097	-1.064	-1.077
	-1.282	-0.899	-0.902	-0.888	-0.882	-0.872	-0.882	-0.857	-0.869	-0.843	-0.855	-0.829	-0.839
Lithuania													
100	-2.326	-1.398	-1.567	-1.169	-1.524	-0.776	-1.471	-0.695	-1.494	-0.605	-1.384	-0.505	-1.364
	-2.054	-1.299	-1.422	-1.109	-1.382	-0.857	-1.343	-0.838	-1.346	-0.764	-1.317	-1.064	-1.253
	-1.645	-1.074	-1.159	-0.937	-1.136	-0.727	-1.097	-0.493	-1.053	-0.518	-1.002	-0.458	-1.010
	-1.282	-0.851	-0.912	-0.748	-0.893	-0.598	-0.865	-0.481	-0.852	-0.413	-0.820	-0.386	-0.792
500	-2.326	-1.545	-1.643	-1.390	-1.624	-1.140	-1.588	-0.899	-1.562	-0.917	-1.565	-0.785	-1.537
	-2.054	-1.369	-1.449	-1.246	-1.436	-1.061	-1.412	-0.884	-1.380	-0.633	-1.356	-0.478	-1.335
	-1.645	-1.099	-1.162	-1.008	-1.152	-0.861	-1.135	-0.722	-1.113	-0.578	-1.096	-0.531	-1.071
	-1.282	-0.854	-0.903	-0.790	-0.896	-0.683	-0.883	-0.568	-0.867	-0.466	-0.855	-0.418	-0.831
1000	-2.326	-1.550	-1.640	-1.417	-1.625	-1.205	-1.599	-1.055	-1.574	-0.875	-1.538	-0.744	-1.518
	-2.054	-1.371	-1.448	-1.268	-1.437	-1.084	-1.417	-0.936	-1.388	-0.823	-1.355	-0.681	-1.335
	-1.645	-1.097	-1.158	-1.015	-1.150	-0.890	-1.133	-0.749	-1.115	-0.628	-1.095	-0.516	-1.072
	-1.282	-0.856	-0.903	-0.795	-0.897	-0.701	-0.884	-0.606	-0.870	-0.508	-0.856	-0.447	-0.838
5000	-2.326	-1.552	-1.637	-1.441	-1.624	-1.283	-1.602	-1.111	-1.578	-0.984	-1.551	-0.848	-1.526
	-2.054	-1.371	-1.446	-1.273	-1.435	-1.142	-1.415	-1.004	-1.393	-0.886	-1.369	-0.767	-1.342
	-1.645	-1.098	-1.158	-1.020	-1.149	-0.914	-1.133	-0.816	-1.116	-0.718	-1.099	-0.630	-1.079
	-1.282	-0.856	-0.902	-0.795	-0.895	-0.713	-0.883	-0.642	-0.870	-0.564	-0.855	-0.488	-0.840

Table 8: Estimates of parameter α : Scenario 1.

Portfolio size	ρ	0.02		0.05		0.10		0.15		0.20		0.25	
	B	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR
Germany													
100	-2.326	0.295	0.714	0.393	0.733	0.460	0.740	0.511	0.775	0.557	0.658	0.575	0.666
	-2.054	0.306	0.697	0.391	0.704	0.454	0.722	0.494	0.737	0.518	0.738	0.538	0.750
	-1.645	0.332	0.690	0.396	0.692	0.444	0.702	0.479	0.709	0.508	0.716	0.533	0.725
	-1.282	0.355	0.701	0.410	0.698	0.443	0.698	0.468	0.702	0.494	0.709	0.511	0.714
500	-2.326	0.396	0.698	0.432	0.697	0.460	0.699	0.488	0.704	0.510	0.708	0.523	0.707
	-2.054	0.410	0.702	0.439	0.700	0.459	0.700	0.477	0.701	0.497	0.704	0.512	0.706
	-1.645	0.424	0.705	0.448	0.704	0.463	0.703	0.475	0.703	0.489	0.703	0.503	0.705
	-1.282	0.432	0.706	0.452	0.705	0.465	0.705	0.476	0.704	0.486	0.704	0.497	0.705
1000	-2.326	0.424	0.702	0.447	0.701	0.463	0.700	0.479	0.701	0.500	0.704	0.518	0.707
	-2.054	0.434	0.705	0.452	0.704	0.464	0.703	0.477	0.703	0.491	0.703	0.508	0.704
	-1.645	0.441	0.705	0.456	0.706	0.468	0.705	0.477	0.705	0.487	0.705	0.498	0.705
	-1.282	0.445	0.707	0.459	0.707	0.469	0.707	0.478	0.706	0.487	0.706	0.496	0.706
5000	-2.326	0.452	0.706	0.461	0.706	0.470	0.705	0.479	0.705	0.488	0.704	0.500	0.704
	-2.054	0.454	0.706	0.462	0.706	0.471	0.706	0.479	0.706	0.487	0.705	0.497	0.705
	-1.645	0.456	0.707	0.463	0.707	0.472	0.707	0.480	0.707	0.488	0.706	0.496	0.706
	-1.282	0.456	0.707	0.463	0.707	0.472	0.707	0.480	0.707	0.488	0.707	0.495	0.707
Lithuania													
100	-2.326	0.210	0.742	0.252	0.761	0.275	0.738	0.306	0.705	0.332	0.723	0.346	0.711
	-2.054	0.209	0.713	0.256	0.720	0.344	0.770	0.393	0.786	0.413	0.796	0.505	0.859
	-1.645	0.203	0.699	0.246	0.712	0.278	0.724	0.317	0.725	0.361	0.740	0.392	0.751
	-1.282	0.202	0.696	0.235	0.703	0.278	0.716	0.314	0.717	0.348	0.724	0.379	0.725
500	-2.326	0.209	0.696	0.243	0.703	0.285	0.713	0.316	0.708	0.357	0.702	0.382	0.704
	-2.054	0.210	0.699	0.238	0.703	0.279	0.708	0.308	0.708	0.331	0.711	0.346	0.712
	-1.645	0.211	0.703	0.234	0.703	0.270	0.706	0.300	0.709	0.327	0.709	0.353	0.710
	-1.282	0.213	0.704	0.232	0.704	0.265	0.704	0.294	0.706	0.319	0.706	0.347	0.709
1000	-2.326	0.212	0.701	0.239	0.702	0.275	0.705	0.315	0.707	0.331	0.715	0.350	0.715
	-2.054	0.213	0.703	0.236	0.703	0.271	0.706	0.304	0.712	0.337	0.715	0.361	0.710
	-1.645	0.214	0.704	0.233	0.704	0.267	0.705	0.294	0.707	0.317	0.708	0.343	0.708
	-1.282	0.214	0.705	0.231	0.705	0.261	0.705	0.290	0.706	0.315	0.706	0.340	0.708
5000	-2.326	0.215	0.705	0.233	0.705	0.266	0.706	0.294	0.706	0.320	0.707	0.344	0.706
	-2.054	0.215	0.706	0.232	0.705	0.264	0.705	0.292	0.706	0.317	0.707	0.339	0.708
	-1.645	0.216	0.706	0.232	0.706	0.260	0.706	0.287	0.706	0.312	0.706	0.335	0.706
	-1.282	0.216	0.707	0.232	0.706	0.258	0.706	0.284	0.706	0.308	0.707	0.330	0.706

Note: true value of α is 0.5.

Table 9: Observed default frequency: Scenario 1.

Portfolio size	TTC \ ρ	0.02	0.05	0.10	0.15	0.20	0.25
		Germany					
100	1%	5.1%	5.3%	5.6%	6%	6.3%	6.7%
	2%	7.5%	7.8%	8.2%	8.6%	9%	9.4%
	5%	12.5%	12.8%	13.2%	13.7%	14.1%	14.6%
	10%	18.4%	18.7%	19.2%	19.6%	20.1%	20.5%
500	1%	5.2%	5.4%	5.7%	6.1%	6.4%	6.8%
	2%	7.5%	7.7%	8.2%	8.6%	9%	9.3%
	5%	12.5%	12.8%	13.2%	13.7%	14.2%	14.6%
	10%	18.4%	18.7%	19.2%	19.7%	20.1%	20.5%
1000	1%	5.2%	5.4%	5.7%	6.1%	6.4%	6.8%
	2%	7.5%	7.8%	8.2%	8.6%	9%	9.4%
	5%	12.5%	12.8%	13.2%	13.7%	14.1%	14.6%
	10%	18.5%	18.8%	19.2%	19.7%	20.1%	20.5%
5000	1%	5.1%	5.4%	5.7%	6.1%	6.4%	6.8%
	2%	7.5%	7.8%	8.2%	8.6%	9%	9.4%
	5%	12.5%	12.8%	13.2%	13.7%	14.1%	14.6%
	10%	18.5%	18.8%	19.2%	19.7%	20.1%	20.5%
Lithuania							
100	1%	6.1%	7.7%	10.1%	12.1%	13.9%	15.5%
	2%	8.6%	10.4%	12.9%	15.1%	16.9%	18.4%
	5%	13.8%	15.7%	18.3%	20.4%	22.1%	23.6%
	10%	19.7%	21.6%	24%	25.9%	27.4%	28.7%
500	1%	6.1%	7.7%	10%	12.1%	13.9%	15.4%
	2%	8.6%	10.4%	12.9%	15.1%	16.9%	18.4%
	5%	13.7%	15.6%	18.2%	20.3%	22%	23.5%
	10%	19.8%	21.6%	24.1%	26%	27.5%	28.7%
1000	1%	6.1%	7.7%	10.1%	12.1%	13.9%	15.4%
	2%	8.6%	10.4%	12.9%	15%	16.9%	18.4%
	5%	13.7%	15.7%	18.3%	20.4%	22.1%	23.5%
	10%	19.7%	21.6%	24%	25.9%	27.4%	28.7%
5000	1%	6.1%	7.7%	10%	12.1%	13.9%	15.4%
	2%	8.6%	10.4%	12.9%	15%	16.9%	18.4%
	5%	13.7%	15.6%	18.3%	20.4%	22.1%	23.5%
	10%	19.7%	21.6%	24%	25.9%	27.4%	28.7%

Table 10: Estimates of parameter ρ by hybrid method (average of GMM and TSR): Scenario 1.

Portfolio size	ρ	0.02	0.05	0.10	0.15	0.20	0.25
	B						
Germany							
100	-2.326	0.034	0.054	0.079	0.100	0.133	0.171
	-2.054	0.033	0.055	0.087	0.113	0.137	0.152
	-1.645	0.029	0.055	0.092	0.122	0.152	0.180
	-1.282	0.026	0.052	0.092	0.127	0.157	0.185
500	-2.326	0.022	0.048	0.087	0.120	0.148	0.176
	-2.054	0.021	0.047	0.088	0.124	0.157	0.185
	-1.645	0.020	0.045	0.086	0.125	0.159	0.190
	-1.282	0.019	0.045	0.086	0.124	0.160	0.194
1000	-2.326	0.020	0.045	0.087	0.123	0.155	0.184
	-2.054	0.019	0.045	0.086	0.124	0.157	0.189
	-1.645	0.019	0.044	0.085	0.124	0.160	0.193
	-1.282	0.018	0.044	0.085	0.123	0.160	0.195
5000	-2.326	0.018	0.043	0.084	0.123	0.159	0.192
	-2.054	0.018	0.043	0.084	0.123	0.159	0.193
	-1.645	0.018	0.043	0.084	0.122	0.159	0.194
	-1.282	0.018	0.043	0.084	0.122	0.159	0.194
Lithuania							
100	-2.326	0.062	0.109	0.172	0.186	0.227	-
	-2.054	0.065	0.111	0.181	0.205	0.198	0.237
	-1.645	0.069	0.123	0.186	0.244	0.304	0.269
	-1.282	0.070	0.130	0.193	0.235	0.256	0.271
500	-2.326	0.066	0.122	0.183	0.226	0.282	0.301
	-2.054	0.065	0.127	0.192	0.240	0.279	0.314
	-1.645	0.065	0.133	0.202	0.251	0.292	0.318
	-1.282	0.064	0.135	0.209	0.262	0.302	0.327
1000	-2.326	0.065	0.128	0.193	0.236	0.263	0.312
	-2.054	0.065	0.131	0.199	0.252	0.282	0.314
	-1.645	0.064	0.134	0.209	0.259	0.302	0.333
	-1.282	0.063	0.135	0.213	0.265	0.305	0.335
5000	-2.326	0.063	0.133	0.206	0.259	0.299	0.333
	-2.054	0.063	0.135	0.211	0.263	0.306	0.342
	-1.645	0.063	0.135	0.216	0.270	0.313	0.345
	-1.282	0.063	0.135	0.219	0.276	0.317	0.352

Table 11: Estimates of parameter B by hybrid method (average of GMM and TSR): Scenario 1.

Portfolio size	ρ B	0.02	0.05	0.10	0.15	0.20	0.25
Germany							
100	-2.326	-1.587	-1.533	-1.462	-1.422	-1.398	-1.352
	-2.054	-1.442	-1.404	-1.350	-1.295	-1.240	-1.217
	-1.645	-1.170	-1.154	-1.115	-1.079	-1.034	-0.988
	-1.282	-0.906	-0.897	-0.878	-0.854	-0.826	-0.798
500	-2.326	-1.638	-1.621	-1.591	-1.554	-1.511	-1.465
	-2.054	-1.445	-1.431	-1.407	-1.377	-1.346	-1.316
	-1.645	-1.157	-1.146	-1.127	-1.107	-1.085	-1.061
	-1.282	-0.901	-0.892	-0.878	-0.864	-0.848	-0.832
1000	-2.326	-1.637	-1.622	-1.595	-1.564	-1.531	-1.492
	-2.054	-1.443	-1.429	-1.406	-1.381	-1.355	-1.322
	-1.645	-1.157	-1.146	-1.127	-1.109	-1.089	-1.067
	-1.282	-0.901	-0.892	-0.878	-0.863	-0.849	-0.834
5000	-2.326	-1.635	-1.619	-1.593	-1.567	-1.541	-1.512
	-2.054	-1.443	-1.429	-1.406	-1.384	-1.361	-1.336
	-1.645	-1.156	-1.145	-1.127	-1.108	-1.090	-1.072
	-1.282	-0.900	-0.892	-0.877	-0.863	-0.849	-0.835
Lithuania							
100	-2.326	-1.503	-1.399	-1.272	-1.173	-1.077	-
	-2.054	-1.353	-1.241	-1.082	-1.012	-0.905	-0.858
	-1.645	-1.118	-1.037	-0.908	-0.792	-0.744	-0.657
	-1.282	-0.887	-0.823	-0.731	-0.656	-0.614	-0.547
500	-2.326	-1.594	-1.510	-1.371	-1.256	-1.124	-1.077
	-2.054	-1.410	-1.341	-1.228	-1.126	-1.037	-0.974
	-1.645	-1.127	-1.079	-0.996	-0.914	-0.844	-0.804
	-1.282	-0.880	-0.845	-0.785	-0.721	-0.659	-0.618
1000	-2.326	-1.598	-1.526	-1.406	-1.320	-1.252	-1.114
	-2.054	-1.410	-1.350	-1.251	-1.148	-1.077	-0.977
	-1.645	-1.129	-1.085	-1.012	-0.939	-0.864	-0.805
	-1.282	-0.880	-0.845	-0.792	-0.735	-0.679	-0.633
5000	-2.326	-1.595	-1.532	-1.445	-1.350	-1.262	-1.188
	-2.054	-1.409	-1.355	-1.279	-1.204	-1.120	-1.040
	-1.645	-1.127	-1.083	-1.023	-0.967	-0.905	-0.853
	-1.282	-0.878	-0.844	-0.798	-0.754	-0.709	-0.662

Table 12: Estimates of parameter α by hybrid method (average of GMM and TSR): Scenario 1.

Portfolio size	ρ B	0.02	0.05	0.10	0.15	0.20	0.25
Germany							
100	-2.326	0.387	0.488	0.568	0.619	0.621	0.604
	-2.054	0.395	0.482	0.542	0.583	0.613	0.648
	-1.645	0.423	0.484	0.527	0.558	0.577	0.594
	-1.282	0.448	0.499	0.529	0.549	0.570	0.586
500	-2.326	0.486	0.517	0.542	0.565	0.586	0.600
	-2.054	0.499	0.524	0.540	0.555	0.569	0.585
	-1.645	0.513	0.532	0.545	0.554	0.565	0.577
	-1.282	0.518	0.535	0.546	0.555	0.563	0.572
1000	-2.326	0.514	0.532	0.544	0.557	0.572	0.586
	-2.054	0.520	0.536	0.546	0.556	0.568	0.579
	-1.645	0.527	0.539	0.549	0.556	0.564	0.574
	-1.282	0.530	0.541	0.550	0.557	0.564	0.571
5000	-2.326	0.536	0.543	0.551	0.558	0.566	0.575
	-2.054	0.537	0.544	0.551	0.558	0.565	0.573
	-1.645	0.539	0.545	0.552	0.559	0.565	0.572
	-1.282	0.539	0.545	0.552	0.559	0.565	0.572
Lithuania							
100	-2.326	0.290	0.342	0.386	0.421	0.439	—
	-2.054	0.282	0.340	0.376	0.434	0.502	0.513
	-1.645	0.275	0.324	0.372	0.393	0.408	0.435
	-1.282	0.274	0.315	0.364	0.404	0.448	0.481
500	-2.326	0.283	0.326	0.375	0.413	0.427	0.463
	-2.054	0.284	0.320	0.366	0.400	0.428	0.449
	-1.645	0.285	0.313	0.356	0.390	0.416	0.446
	-1.282	0.287	0.310	0.350	0.382	0.410	0.439
1000	-2.326	0.286	0.319	0.365	0.402	0.440	0.451
	-2.054	0.287	0.315	0.358	0.389	0.424	0.450
	-1.645	0.288	0.311	0.350	0.383	0.409	0.435
	-1.282	0.289	0.311	0.347	0.379	0.407	0.434
5000	-2.326	0.290	0.313	0.353	0.384	0.412	0.436
	-2.054	0.290	0.311	0.348	0.381	0.407	0.430
	-1.645	0.291	0.311	0.344	0.375	0.402	0.428
	-1.282	0.291	0.311	0.342	0.372	0.399	0.423

Note: true value of α is 0.5.

Table 13: Estimates of Vasicek correlation parameter ρ : Scenario 2.

Portfolio size	ρ	0.02	0.05	0.10	0.15	0.20	0.25
	B	GMM TSR	GMM TSR	GMM TSR	GMM TSR	GMM TSR	GMM TSR
Germany							
100	-2.326	0.062 0.012 (0.434)	0.091 0.027 (0.734)	0.131 0.051 (0.943)	0.162 0.077 (0.979)	0.197 0.098 (0.993)	0.274 0.135 (0.999)
	-2.054	0.055 0.012 (0.098)	0.088 0.028 (0.357)	0.133 0.054 (0.709)	0.168 0.077 (0.885)	0.196 0.101 (0.959)	0.232 0.126 (0.983)
	-1.645	0.046 0.011 (0.002)	0.084 0.027 (0.028)	0.136 0.053 (0.264)	0.175 0.078 (0.552)	0.208 0.102 (0.768)	0.231 0.124 (0.881)
	-1.282	0.040 0.010 (0.000)	0.076 0.026 (0.002)	0.131 0.051 (0.044)	0.177 0.076 (0.179)	0.216 0.100 (0.392)	0.245 0.123 (0.600)
500	-2.326	0.035 0.011 (0.000)	0.075 0.028 (0.006)	0.132 0.056 (0.169)	0.179 0.083 (0.446)	0.214 0.110 (0.689)	0.247 0.134 (0.841)
	-2.054	0.032 0.011 (0.000)	0.071 0.027 (0.006)	0.131 0.055 (0.169)	0.180 0.082 (0.446)	0.219 0.108 (0.689)	0.254 0.134 (0.841)
	-1.645	0.029 0.010 (0.000)	0.067 0.026 (0.000)	0.125 0.052 (0.002)	0.178 0.079 (0.033)	0.223 0.105 (0.145)	0.260 0.130 (0.308)
	-1.282	0.027 0.010 (0.000)	0.063 0.025 (0.000)	0.118 0.049 (0.001)	0.170 0.074 (0.004)	0.216 0.099 (0.030)	0.257 0.124 (0.102)
1000	-2.326	0.031 0.011 (0.000)	0.070 0.028 (0.000)	0.131 0.055 (0.035)	0.182 0.083 (0.193)	0.223 0.110 (0.418)	0.256 0.136 (0.648)
	-2.054	0.029 0.011 (0.000)	0.068 0.027 (0.000)	0.128 0.054 (0.035)	0.182 0.082 (0.193)	0.227 0.109 (0.418)	0.267 0.135 (0.648)
	-1.645	0.028 0.010 (0.000)	0.065 0.026 (0.000)	0.123 0.052 (0.000)	0.177 0.078 (0.005)	0.225 0.104 (0.049)	0.266 0.130 (0.156)
	-1.282	0.026 0.010 (0.000)	0.061 0.025 (0.000)	0.117 0.049 (0.000)	0.168 0.074 (0.001)	0.217 0.099 (0.002)	0.260 0.124 (0.033)
5000	-2.326	0.028 0.011 (0.000)	0.067 0.027 (0.000)	0.128 0.055 (0.002)	0.184 0.083 (0.011)	0.233 0.110 (0.069)	0.276 0.138 (0.193)
	-2.054	0.027 0.011 (0.000)	0.065 0.027 (0.000)	0.125 0.054 (0.000)	0.181 0.081 (0.002)	0.231 0.108 (0.021)	0.276 0.135 (0.075)
	-1.645	0.026 0.010 (0.000)	0.063 0.026 (0.000)	0.121 0.052 (0.000)	0.175 0.078 (0.000)	0.225 0.104 (0.005)	0.271 0.130 (0.017)
	-1.282	0.025 0.010 (0.000)	0.060 0.024 (0.000)	0.115 0.049 (0.000)	0.167 0.074 (0.000)	0.215 0.099 (0.002)	0.260 0.123 (0.003)
Lithuania							
100	-2.326	0.124 0.011 (0.903)	0.205 0.024 (0.990)	0.313 0.034 (0.999)	0.413 0.052 (0.999)	0.498 0.072 (0.999)	— — (1.000)
Continued on next page							

Table 13 – continued from previous page

Portfolio size	ρ	0.02	0.05	0.10	0.15	0.20	0.25
	B	GMM TSR	GMM TSR	GMM TSR	GMM TSR	GMM TSR	GMM TSR
100	-2.054	0.128 0.011 (0.650)	0.207 0.025 (0.934)	0.308 0.046 (0.992)	0.379 0.069 (0.995)	0.448 0.093 (0.999)	– – (1.000)
	-1.645	0.131 0.011 (0.229)	0.226 0.025 (0.696)	0.314 0.048 (0.945)	0.350 0.067 (0.990)	0.398 0.085 (0.994)	0.369 0.106 (0.997)
	-1.282	0.126 0.010 (0.050)	0.227 0.025 (0.429)	0.326 0.046 (0.823)	0.396 0.067 (0.941)	0.413 0.085 (0.977)	0.434 0.106 (0.991)
500	-2.326	0.128 0.011 (0.121)	0.229 0.027 (0.635)	0.325 0.052 (0.932)	0.393 0.076 (0.985)	0.442 0.097 (0.993)	0.438 0.125 (0.996)
	-2.054	0.127 0.011 (0.037)	0.236 0.027 (0.434)	0.336 0.051 (0.849)	0.400 0.074 (0.954)	0.436 0.093 (0.984)	0.465 0.111 (0.995)
	-1.645	0.122 0.010 (0.003)	0.241 0.026 (0.171)	0.352 0.050 (0.636)	0.421 0.072 (0.851)	0.464 0.093 (0.946)	0.489 0.113 (0.978)
	-1.282	0.116 0.010 (0.000)	0.239 0.024 (0.048)	0.355 0.048 (0.392)	0.428 0.069 (0.704)	0.478 0.089 (0.854)	0.509 0.109 (0.935)
1000	-2.326	0.127 0.011 (0.019)	0.240 0.027 (0.408)	0.348 0.053 (0.832)	0.410 0.075 (0.957)	0.424 0.096 (0.987)	0.462 0.117 (0.993)
	-2.054	0.124 0.011 (0.009)	0.244 0.027 (0.231)	0.354 0.051 (0.701)	0.428 0.075 (0.899)	0.461 0.097 (0.966)	0.497 0.119 (0.983)
	-1.645	0.120 0.010 (0.000)	0.246 0.026 (0.066)	0.364 0.050 (0.450)	0.436 0.073 (0.726)	0.490 0.094 (0.883)	0.513 0.115 (0.949)
	-1.282	0.114 0.010 (0.000)	0.239 0.024 (0.020)	0.365 0.048 (0.247)	0.439 0.070 (0.559)	0.491 0.091 (0.758)	0.525 0.110 (0.887)
5000	-2.326	0.125 0.011 (0.000)	0.254 0.027 (0.082)	0.372 0.053 (0.494)	0.448 0.077 (0.764)	0.501 0.100 (0.901)	0.535 0.122 (0.955)
	-2.054	0.122 0.011 (0.000)	0.252 0.026 (0.039)	0.375 0.052 (0.347)	0.452 0.076 (0.651)	0.504 0.099 (0.825)	0.538 0.121 (0.910)
	-1.645	0.118 0.010 (0.000)	0.248 0.026 (0.039)	0.382 0.050 (0.347)	0.460 0.074 (0.651)	0.513 0.097 (0.825)	0.552 0.118 (0.910)
	-1.282	0.113 0.010 (0.000)	0.239 0.024 (0.000)	0.377 0.048 (0.078)	0.461 0.071 (0.276)	0.513 0.092 (0.526)	0.557 0.114 (0.696)

Note: in the parentheses below estimates is specified the ratio of simulations in which at at least one point in time t the number of defaults was equal to zero.

Table 14: Estimates of parameter B : Scenario 2.

Portfolio size	ρ	0.02		0.05		0.10		0.15		0.20		0.25	
	B	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR
Germany													
100	-2.326	-1.502	-1.543	-1.450	-1.512	-1.365	-1.481	-1.291	-1.436	-1.220	-1.383	-1.059	-1.495
	-2.054	-1.328	-1.355	-1.288	-1.328	-1.230	-1.300	-1.170	-1.262	-1.106	-1.234	-1.047	-1.219
	-1.645	-1.058	-1.073	-1.044	-1.065	-1.010	-1.045	-0.968	-1.020	-0.927	-0.999	-0.892	-0.976
	-1.282	-0.788	-0.797	-0.781	-0.792	-0.766	-0.782	-0.744	-0.768	-0.715	-0.753	-0.679	-0.731
500	-2.326	-1.537	-1.549	-1.517	-1.539	-1.479	-1.519	-1.434	-1.497	-1.390	-1.471	-1.343	-1.445
	-2.054	-1.341	-1.349	-1.325	-1.340	-1.298	-1.325	-1.266	-1.307	-1.230	-1.286	-1.190	-1.263
	-1.645	-1.053	-1.058	-1.041	-1.051	-1.023	-1.039	-1.004	-1.025	-0.982	-1.010	-0.954	-0.992
	-1.282	-0.790	-0.793	-0.782	-0.788	-0.770	-0.779	-0.759	-0.769	-0.748	-0.759	-0.732	-0.747
1000	-2.326	-1.535	-1.544	-1.516	-1.534	-1.485	-1.516	-1.453	-1.496	-1.417	-1.473	-1.372	-1.447
	-2.054	-1.339	-1.346	-1.324	-1.337	-1.299	-1.322	-1.273	-1.305	-1.247	-1.287	-1.213	-1.266
	-1.645	-1.054	-1.058	-1.042	-1.051	-1.024	-1.039	-1.007	-1.025	-0.988	-1.011	-0.968	-0.995
	-1.282	-0.790	-0.792	-0.782	-0.787	-0.770	-0.778	-0.759	-0.768	-0.749	-0.758	-0.737	-0.747
5000	-2.326	-1.536	-1.542	-1.517	-1.532	-1.489	-1.514	-1.462	-1.494	-1.434	-1.473	-1.403	-1.450
	-2.054	-1.339	-1.345	-1.324	-1.336	-1.300	-1.320	-1.277	-1.303	-1.255	-1.284	-1.231	-1.265
	-1.645	-1.053	-1.057	-1.042	-1.050	-1.024	-1.037	-1.007	-1.024	-0.992	-1.009	-0.977	-0.994
	-1.282	-0.789	-0.791	-0.781	-0.786	-0.769	-0.777	-0.758	-0.767	-0.748	-0.756	-0.739	-0.745
Lithuania													
100	-2.326	-1.325	-1.508	-1.087	-1.502	-0.552	-1.388	-0.382	-1.378	-0.255	-1.380	-	-
	-2.054	-1.215	-1.338	-1.050	-1.315	-0.744	-1.305	-0.496	-1.275	-0.607	-1.372	-	-
	-1.645	-0.985	-1.065	-0.858	-1.049	-0.659	-1.012	-0.436	-0.994	-0.305	-0.996	-0.484	-1.014
	-1.282	-0.751	-0.800	-0.676	-0.792	-0.518	-0.776	-0.404	-0.776	-0.318	-0.750	-0.262	-0.717
500	-2.326	-1.455	-1.553	-1.316	-1.543	-1.097	-1.523	-0.881	-1.515	-0.680	-1.475	-0.724	-1.450
	-2.054	-1.273	-1.353	-1.157	-1.345	-0.973	-1.326	-0.800	-1.308	-0.687	-1.265	-0.467	-1.232
	-1.645	-1.006	-1.064	-0.930	-1.061	-0.790	-1.050	-0.667	-1.036	-0.541	-1.018	-0.418	-0.991
	-1.282	-0.757	-0.798	-0.711	-0.800	-0.626	-0.797	-0.514	-0.792	-0.421	-0.783	-0.335	-0.770
1000	-2.326	-1.459	-1.549	-1.334	-1.542	-1.138	-1.524	-0.935	-1.496	-0.840	-1.451	-0.715	-1.414
	-2.054	-1.275	-1.351	-1.179	-1.346	-1.013	-1.332	-0.844	-1.315	-0.717	-1.282	-0.648	-1.252
	-1.645	-1.005	-1.061	-0.935	-1.060	-0.822	-1.053	-0.708	-1.043	-0.594	-1.030	-0.513	-1.012
	-1.282	-0.756	-0.796	-0.710	-0.797	-0.638	-0.797	-0.553	-0.792	-0.462	-0.785	-0.368	-0.772
5000	-2.326	-1.459	-1.545	-1.352	-1.539	-1.202	-1.525	-1.055	-1.507	-0.911	-1.484	-0.787	-1.456
	-2.054	-1.275	-1.348	-1.187	-1.344	-1.066	-1.334	-0.941	-1.320	-0.822	-1.302	-0.738	-1.279
	-1.645	-1.005	-1.061	-0.940	-1.059	-0.850	-1.055	-0.768	-1.048	-0.681	-1.037	-0.576	-1.022
	-1.282	-0.756	-0.795	-0.712	-0.797	-0.652	-0.798	-0.593	-0.796	-0.526	-0.791	-0.455	-0.783

Table 15: Estimates of parameter α : Scenario 2.

Portfolio size	ρ	0.02		0.05		0.10		0.15		0.20		0.25	
	B	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR	GMM	TSR
Germany													
100	-2.326	0.288	0.668	0.375	0.684	0.443	0.703	0.490	0.704	0.515	0.714	0.478	0.681
	-2.054	0.307	0.672	0.382	0.677	0.440	0.687	0.479	0.702	0.512	0.707	0.525	0.706
	-1.645	0.337	0.687	0.395	0.682	0.438	0.688	0.471	0.696	0.498	0.702	0.527	0.712
	-1.282	0.361	0.703	0.415	0.701	0.446	0.702	0.468	0.705	0.489	0.709	0.511	0.716
500	-2.326	0.388	0.671	0.418	0.668	0.444	0.669	0.466	0.672	0.490	0.676	0.510	0.683
	-2.054	0.405	0.683	0.430	0.680	0.446	0.678	0.464	0.679	0.485	0.681	0.502	0.685
	-1.645	0.423	0.698	0.444	0.695	0.457	0.693	0.468	0.692	0.482	0.692	0.497	0.694
	-1.282	0.439	0.716	0.458	0.714	0.471	0.713	0.480	0.712	0.490	0.711	0.500	0.712
1000	-2.326	0.413	0.676	0.432	0.674	0.446	0.672	0.462	0.673	0.481	0.675	0.500	0.678
	-2.054	0.423	0.683	0.439	0.682	0.451	0.680	0.463	0.679	0.476	0.679	0.490	0.681
	-1.645	0.437	0.696	0.451	0.695	0.461	0.694	0.470	0.693	0.479	0.693	0.492	0.693
	-1.282	0.451	0.715	0.463	0.714	0.473	0.713	0.482	0.712	0.489	0.711	0.498	0.711
5000	-2.326	0.435	0.677	0.443	0.676	0.452	0.675	0.460	0.674	0.470	0.673	0.483	0.674
	-2.054	0.441	0.685	0.448	0.684	0.457	0.683	0.465	0.682	0.473	0.681	0.482	0.681
	-1.645	0.449	0.696	0.456	0.696	0.464	0.695	0.472	0.694	0.480	0.694	0.488	0.693
	-1.282	0.462	0.715	0.468	0.715	0.476	0.714	0.484	0.713	0.491	0.713	0.498	0.712
Lithuania													
100	-2.326	0.204	0.692	0.248	0.723	0.283	0.854	0.302	0.846	0.317	0.831	—	—
	-2.054	0.202	0.683	0.250	0.711	0.291	0.739	0.320	0.739	0.334	0.734	—	—
	-1.645	0.200	0.688	0.240	0.705	0.286	0.721	0.335	0.751	0.363	0.768	0.417	0.767
	-1.282	0.205	0.707	0.240	0.717	0.281	0.741	0.312	0.752	0.351	0.773	0.382	0.769
500	-2.326	0.203	0.670	0.238	0.679	0.282	0.695	0.313	0.704	0.342	0.720	0.381	0.711
	-2.054	0.205	0.679	0.235	0.687	0.277	0.702	0.310	0.711	0.343	0.732	0.371	0.750
	-1.645	0.210	0.694	0.233	0.698	0.270	0.711	0.301	0.724	0.332	0.734	0.361	0.746
	-1.282	0.215	0.714	0.235	0.716	0.269	0.726	0.299	0.739	0.326	0.750	0.353	0.760
1000	-2.326	0.205	0.675	0.233	0.680	0.272	0.691	0.307	0.708	0.346	0.723	0.370	0.735
	-2.054	0.208	0.683	0.232	0.686	0.269	0.697	0.299	0.708	0.332	0.719	0.358	0.727
	-1.645	0.211	0.696	0.231	0.698	0.266	0.708	0.296	0.718	0.322	0.730	0.351	0.737
	-1.282	0.217	0.716	0.235	0.718	0.266	0.724	0.295	0.735	0.322	0.744	0.347	0.755
5000	-2.326	0.208	0.678	0.227	0.680	0.263	0.688	0.292	0.696	0.319	0.706	0.345	0.717
	-2.054	0.210	0.686	0.228	0.688	0.261	0.694	0.291	0.703	0.318	0.712	0.343	0.721
	-1.645	0.213	0.698	0.230	0.700	0.260	0.705	0.288	0.713	0.314	0.720	0.338	0.729
	-1.282	0.219	0.717	0.235	0.719	0.262	0.723	0.288	0.729	0.315	0.737	0.337	0.743

Note: true value of α is 0.5.

Table 16: Observed default frequency: Scenario 2.

Portfolio size	ρ	0.02	0.05	0.10	0.15	0.20	0.25
	TTC						
Germany							
100	1%	6.3%	6.5%	6.9%	7.3%	7.7%	8.1%
	2%	9.1%	9.4%	9.9%	10.3%	10.7%	11.1%
	5%	14.6%	14.9%	15.3%	15.7%	16.2%	16.6%
	10%	21.6%	21.9%	22.3%	22.6%	23.0%	23.3%
500	1%	6.2%	6.5%	6.9%	7.3%	7.7%	8.0%
	2%	9.0%	9.3%	9.8%	10.2%	10.6%	11.0%
	5%	14.7%	14.9%	15.4%	15.8%	16.3%	16.7%
	10%	21.5%	21.8%	22.2%	22.5%	22.9%	23.2%
1000	1%	6.3%	6.5%	6.9%	7.3%	7.7%	8.0%
	2%	9.0%	9.3%	9.8%	10.2%	10.6%	11.0%
	5%	14.6%	14.9%	15.4%	15.8%	16.2%	16.6%
	10%	21.5%	21.8%	22.2%	22.5%	22.9%	23.2%
5000	1%	6.2%	6.5%	6.9%	7.3%	7.7%	8.0%
	2%	9.0%	9.3%	9.8%	10.2%	10.6%	11.0%
	5%	14.6%	14.9%	15.4%	15.8%	16.2%	16.6%
	10%	21.5%	21.8%	22.2%	22.5%	22.9%	23.2%
Lithuania							
100	1%	7.3%	9.0%	11.4%	13.4%	15.1%	16.6%
	2%	10.2%	12.0%	14.4%	16.4%	18.1%	19.5%
	5%	15.9%	17.7%	20.0%	21.9%	23.4%	24.7%
	10%	22.7%	24.2%	26.2%	27.7%	29.0%	30.0%
500	1%	7.3%	9.0%	11.3%	13.3%	15.0%	16.5%
	2%	10.2%	12.0%	14.5%	16.5%	18.1%	19.6%
	5%	15.9%	17.6%	20.0%	21.9%	23.4%	24.7%
	10%	22.6%	24.1%	26.1%	27.6%	28.9%	29.9%
1000	1%	7.3%	9.0%	11.4%	13.3%	15.0%	16.5%
	2%	10.2%	12.0%	14.5%	16.5%	18.1%	19.6%
	5%	15.9%	17.7%	20.0%	21.9%	23.4%	24.7%
	10%	22.6%	24.1%	26.1%	27.6%	28.9%	30.0%
5000	1%	7.3%	9.0%	11.4%	13.4%	15.1%	16.5%
	2%	10.2%	12.0%	14.5%	16.5%	18.1%	19.6%
	5%	15.8%	17.6%	20%	21.9%	23.4%	24.7%
	10%	22.6%	24.1%	26.1%	27.6%	28.9%	30.0%

Table 17: Estimates of parameter ρ by hybrid method (average of GMM and TSR): Scenario 2.

Portfolio size	ρ	0.02	0.05	0.10	0.15	0.20	0.25
	B						
Germany							
100	-2.326	0.036	0.059	0.090	0.115	0.152	0.172
	-2.054	0.034	0.059	0.093	0.123	0.148	0.167
	-1.645	0.028	0.055	0.095	0.128	0.157	0.183
	-1.282	0.025	0.051	0.091	0.126	0.157	0.185
500	-2.326	0.023	0.052	0.095	0.132	0.164	0.191
	-2.054	0.021	0.049	0.093	0.131	0.165	0.195
	-1.645	0.020	0.046	0.089	0.128	0.164	0.196
	-1.282	0.019	0.044	0.084	0.122	0.158	0.191
1000	-2.326	0.021	0.049	0.094	0.133	0.168	0.198
	-2.054	0.020	0.048	0.091	0.131	0.168	0.200
	-1.645	0.019	0.046	0.088	0.127	0.165	0.199
	-1.282	0.018	0.043	0.083	0.121	0.158	0.192
5000	-2.326	0.019	0.047	0.092	0.133	0.172	0.207
	-2.054	0.019	0.046	0.090	0.131	0.169	0.205
	-1.645	0.018	0.045	0.087	0.127	0.165	0.200
	-1.282	0.017	0.042	0.082	0.120	0.157	0.192
Lithuania							
100	-2.326	0.067	0.113	0.181	0.218	0.254	0.286
	-2.054	0.070	0.115	0.170	0.219	0.249	0.298
	-1.645	0.070	0.126	0.187	0.223	0.240	0.274
	-1.282	0.068	0.126	0.188	0.227	0.254	0.275
500	-2.326	0.069	0.129	0.195	0.231	0.267	0.286
	-2.054	0.069	0.132	0.197	0.237	0.261	0.284
	-1.645	0.066	0.133	0.201	0.249	0.279	0.302
	-1.282	0.063	0.131	0.201	0.247	0.284	0.310
1000	-2.326	0.069	0.134	0.200	0.245	0.282	0.293
	-2.054	0.068	0.135	0.201	0.246	0.273	0.298
	-1.645	0.065	0.136	0.206	0.253	0.290	0.314
	-1.282	0.062	0.132	0.206	0.254	0.288	0.317
5000	-2.326	0.068	0.140	0.213	0.262	0.298	0.327
	-2.054	0.066	0.139	0.216	0.264	0.302	0.326
	-1.645	0.064	0.137	0.216	0.267	0.304	0.335
	-1.282	0.061	0.132	0.212	0.265	0.304	0.334

Table 18: Estimates of parameter B by hybrid method (average of GMM and TSR): Scenario 2.

Portfolio size	ρ B	0.02	0.05	0.10	0.15	0.20	0.25
Germany							
100	-2.326	-1.517	-1.471	-1.402	-1.335	-1.257	-1.157
	-2.054	-1.350	-1.317	-1.264	-1.220	-1.166	-1.135
	-1.645	-1.062	-1.049	-1.016	-0.985	-0.953	-0.919
	-1.282	-0.796	-0.790	-0.777	-0.759	-0.738	-0.710
500	-2.326	-1.543	-1.528	-1.498	-1.463	-1.424	-1.388
	-2.054	-1.345	-1.333	-1.312	-1.288	-1.259	-1.225
	-1.645	-1.057	-1.047	-1.032	-1.016	-0.998	-0.978
	-1.282	-0.792	-0.785	-0.775	-0.764	-0.753	-0.741
1000	-2.326	-1.542	-1.527	-1.503	-1.476	-1.447	-1.413
	-2.054	-1.341	-1.329	-1.309	-1.288	-1.264	-1.238
	-1.645	-1.055	-1.046	-1.031	-1.016	-1.000	-0.981
	-1.282	-0.791	-0.785	-0.774	-0.764	-0.753	-0.742
5000	-2.326	-1.539	-1.525	-1.501	-1.478	-1.454	-1.430
	-2.054	-1.342	-1.330	-1.310	-1.290	-1.271	-1.250
	-1.645	-1.055	-1.046	-1.031	-1.016	-1.001	-0.986
	-1.282	-0.789	-0.783	-0.772	-0.762	-0.752	-0.741
Lithuania							
100	-2.326	-1.413	-1.293	-1.131	-1.084	-1.006	-0.938
	-2.054	-1.271	-1.168	-0.972	-0.856	-0.866	-0.890
	-1.645	-1.025	-0.948	-0.833	-0.710	-0.745	-0.706
	-1.282	-0.783	-0.744	-0.658	-0.568	-0.494	-0.434
500	-2.326	-1.502	-1.426	-1.291	-1.211	-1.106	-1.042
	-2.054	-1.312	-1.253	-1.145	-1.060	-0.976	-0.894
	-1.645	-1.034	-0.997	-0.924	-0.841	-0.773	-0.728
	-1.282	-0.777	-0.754	-0.710	-0.659	-0.608	-0.541
1000	-2.326	-1.505	-1.437	-1.325	-1.227	-1.146	-1.130
	-2.054	-1.313	-1.263	-1.169	-1.089	-1.006	-0.932
	-1.645	-1.033	-0.997	-0.938	-0.874	-0.809	-0.755
	-1.282	-0.776	-0.755	-0.718	-0.674	-0.625	-0.579
5000	-2.326	-1.503	-1.446	-1.365	-1.276	-1.199	-1.124
	-2.054	-1.312	-1.266	-1.198	-1.128	-1.068	-0.999
	-1.645	-1.032	-0.999	-0.951	-0.906	-0.856	-0.799
	-1.282	-0.776	-0.755	-0.726	-0.694	-0.661	-0.625

Table 19: Estimates of parameter α by hybrid method (average of GMM and TSR): Scenario 2.

Portfolio size	ρ B	0.02	0.05	0.10	0.15	0.20	0.25
Germany							
100	-2.326	0.375	0.464	0.533	0.575	0.576	0.604
	-2.054	0.392	0.468	0.524	0.557	0.586	0.616
	-1.645	0.426	0.482	0.520	0.546	0.568	0.588
	-1.282	0.456	0.505	0.531	0.550	0.569	0.586
500	-2.326	0.472	0.498	0.519	0.538	0.557	0.575
	-2.054	0.489	0.510	0.525	0.539	0.554	0.569
	-1.645	0.507	0.526	0.537	0.546	0.557	0.569
	-1.282	0.525	0.541	0.552	0.559	0.567	0.576
1000	-2.326	0.495	0.510	0.522	0.535	0.549	0.565
	-2.054	0.505	0.519	0.529	0.539	0.549	0.563
	-1.645	0.519	0.531	0.540	0.547	0.555	0.564
	-1.282	0.536	0.546	0.555	0.562	0.568	0.575
5000	-2.326	0.515	0.521	0.528	0.535	0.543	0.553
	-2.054	0.521	0.527	0.534	0.540	0.547	0.556
	-1.645	0.531	0.536	0.543	0.549	0.556	0.562
	-1.282	0.545	0.551	0.557	0.564	0.569	0.575
Lithuania							
100	-2.326	0.277	0.337	0.372	0.415	0.444	0.468
	-2.054	0.273	0.335	0.389	0.418	0.449	0.458
	-1.645	0.272	0.320	0.369	0.414	0.461	0.481
	-1.282	0.278	0.320	0.369	0.410	0.445	0.478
500	-2.326	0.275	0.316	0.362	0.408	0.437	0.471
	-2.054	0.277	0.312	0.360	0.401	0.441	0.472
	-1.645	0.283	0.312	0.356	0.391	0.426	0.457
	-1.282	0.291	0.315	0.356	0.393	0.422	0.451
1000	-2.326	0.276	0.311	0.357	0.394	0.425	0.464
	-2.054	0.280	0.310	0.356	0.394	0.431	0.461
	-1.645	0.286	0.309	0.352	0.388	0.418	0.448
	-1.282	0.293	0.315	0.353	0.387	0.419	0.446
5000	-2.326	0.280	0.304	0.346	0.381	0.412	0.440
	-2.054	0.283	0.305	0.344	0.379	0.409	0.439
	-1.645	0.288	0.308	0.344	0.377	0.407	0.433
	-1.282	0.295	0.315	0.347	0.379	0.408	0.434

Note: true value of α is 0.5.

8 Appendix B. R code

```
#####  
#                                                                 MAIN  
#####  
  
# packages-----  
packages <- c("tseries", "dplyr", "stargazer", "lubridate", "readxl", "ggplot2",  
             "reshape")  
  
if (length(setdiff(packages, rownames(installed.packages()))) > 0) {  
  install.packages(setdiff(packages, rownames(installed.packages())),  
                  repos = "http://cran.us.r-project.org")  
}  
  
# Load the packages:  
for(i in 1: length(packages)){  
  string_to_eval <- paste("library(", packages[i],")", sep="")  
  eval(parse(text= string_to_eval))  
}  
# -----  
  
# directories and functions-----  
main_dir <- getwd()  
setwd(main_dir)  
  
source(paste(main_dir, "/functions/sourceDir.r", sep=""))  
sourceDir(paste(main_dir, "/functions/", sep=""))  
# -----  
  
# setting parameters and creating objects for simulated data-----  
# parameters for cyclical time series  
DE_param <- list(lambda = .18685, rho = .72067 , rhoc = .33405 , alpha = .00852,  
                sigma_err =.00004864, sigma_level = .00003159, sigma_drift = .00000458)  
LT_param <- list(lambda = .19113, rho = .82559 , rhoc = .56937 , alpha = .01107,  
                sigma_err =.0000404, sigma_level = .0000547, sigma_drift = .0000124)  
  
# number of simulations  
N <- 10000  
  
# lists for macro data  
macro_data_DE <- list()  
macro_data_LT <- list()  
  
# lists for cycles  
cycle_data_DE <- list()  
cycle_data_LT <- list()  
  
# Number of rejected null hypothesis for Germany case  
shapiro_wilk_reject_DE <- numeric(N)  
jarque_bera_reject_DE <- numeric(N)  
# Number of rejected null hypothesis for Lithuania case  
shapiro_wilk_reject_LT <- numeric(N)  
jarque_bera_reject_LT <- numeric(N)  
# -----
```

```

# simulations-----
# generate cycles
for(i in 1:N){
  # Germany
  cyclical_ts_DE <- sim.cyclical.ts(lambda = DE_param$lambda, rho = DE_param$rho,
                                   rhoc = DE_param$rhoc, alfa = DE_param$alpha,
                                   eps_zeta_sd = DE_param$sigma_drift,
                                   eps_eta_sd = DE_param$sigma_level,
                                   eps_irr_sd = DE_param$sigma_err,
                                   m0 = rep(0,4))

  macro_data_DE[[i]] <- cyclical_ts_DE[, "series"]

  cycle_DE <- cyclical_ts_DE[, "cycle"]*100

  # Normality tests
  shapiro_wilk_reject_DE[i] <- ifelse(shapiro.test(cycle_DE)[2] < 0.05, 1, 0)
  jarque_bera_reject_DE[i] <- ifelse(jarque.bera.test(cycle_DE)[3] < 0.05, 1, 0)

  cycle_data_DE[[i]] <- cycle_DE

  # Lithuania
  cyclical_ts_LT <- sim.cyclical.ts(lambda = LT_param$lambda, rho = LT_param$rho,
                                   rhoc = LT_param$rhoc, alfa = LT_param$alpha,
                                   eps_zeta_sd = LT_param$sigma_drift,
                                   eps_eta_sd = LT_param$sigma_level,
                                   eps_irr_sd = LT_param$sigma_err,
                                   m0 = rep(0,4))

  macro_data_LT[[i]] <- cyclical_ts_LT[, "series"]

  cycle_LT <- cyclical_ts_LT[, "cycle"]*100

  # tests
  shapiro_wilk_reject_LT[i] <- ifelse(shapiro.test(cycle_LT)[2] < 0.05, 1, 0)
  jarque_bera_reject_LT[i] <- ifelse(jarque.bera.test(cycle_LT)[3] < 0.05, 1, 0)

  cycle_data_LT[[i]] <- cycle_LT
}
# -----

# aggregated results of normality tests-----
# percent of rejected hypothesis, i.e. non normally distributed Z
sum(shapiro_wilk_reject_DE)/N
sum(jarque_bera_reject_DE)/N

sum(shapiro_wilk_reject_LT)/N
sum(jarque_bera_reject_LT)/N

# Check if long time series is normally distributed
# DE
hist(unlist(cycle_data_DE), col = 'steelblue')

```

```

jarque.bera.test(unlist(cycle_data_DE))

# LT
hist(unlist(cycle_data_LT), col = 'steelblue')
jarque.bera.test(unlist(cycle_data_LT))
# -----

# save generated data-----
saveRDS(matrix(unlist(cycle_data_DE), ncol = 100, byrow = T),
         "../04 Generated Data/cycle_data_DE.RDS")

saveRDS(matrix(unlist(cycle_data_LT), ncol = 100, byrow = T),
         "../04 Generated Data/cycle_data_LT.RDS")

saveRDS(matrix(unlist(macro_data_DE), ncol = 100, byrow = T),
         "../04 Generated Data/macro_data_DE.RDS")

saveRDS(matrix(unlist(macro_data_LT), ncol = 100, byrow = T),
         "../04 Generated Data/macro_data_LT.RDS")
# -----

# controled simulation experiment-----
# parameters
K_max <- 5000
K <- c(100, 500, 1000, 5000)
TTC_PD <- c(0.01, 0.02, 0.05, 0.1)
B_vid <- qnorm(TTC_PD)
rho <- c(0.02, seq(0.05, 0.25, 0.05))
alpha <- 0.5

# for the parameter estimation the function is written. it is used for all estimations
# just country and scenario changes. here one of cases is presented

results_DE_scn1 <- est.TTC_PD.params(country = "DE", scenario = 1,
                                    use_hybrid_method = FALSE)

# save tex tables with mean and CI of estimates
# B, rho and alpha
make_tex_tables(estimates = results$rho_GMM, K, B_vid, rho, N, not_eval = TRUE,
               file_name = "../05 Output/DE/Z_given_B_norm/rho_GMM")

make_tex_tables(estimates = results$B_GMM, K, B_vid, rho, N, not_eval = TRUE,
               file_name = "../05 Output/DE/Z_given_B_norm/B_GMM")

make_tex_tables_alpha(estimates = results$alpha_GMM, K, B_vid, rho, alpha, N,
                    file_name = "../05 Output/DE/Z_given_B_norm/alpha_GMM")

make_tex_tables(estimates = results$rho_TSR, K, B_vid, rho, N,
               file_name = "../05 Output/DE/Z_given_B_norm/rho_TSR")

make_tex_tables(estimates = results$B_TSR, K, B_vid, rho, N,
               file_name = "../05 Output/DE/Z_given_B_norm/B_TSR")

```

```

make_tex_tables_alpha(estimates = results$alpha_TSR, K, B_vid, rho, alpha, N,
                      file_name = "../05 Output/DE/Z_given_B_norm/alpha_TSR")

# ODF
make_tex_tables(estimates = results$ODF, K, B_vid, rho, N,
                file_name = "../05 Output/DE/Z_given_B_norm/ODF")

# -----

# #####
#                                     FUNCTIONS
# #####

# sourceDir-----
## If you want to source() a bunch of files, something like
## the following may be useful:
sourceDir <- function(path, trace = TRUE, ...) {
  for (nm in list.files(path, pattern = "\\.[RrSsQq]$" )) {
    if(trace) cat(nm,":")
    source(file.path(path, nm), ...)
    if(trace) cat("\n")
  }
}

# -----

# sim.cyclical.ts-----
sim.cyclical.ts <- function(n          = 100,
                           lambda     = 0.2,
                           rho        = 0.8,
                           rhoc       = 0.5,
                           alfa       = 0.008,
                           a          = 1,
                           eps_kappa_sd = 1,
                           eps_zeta_sd = 0.0001,
                           eps_eta_sd  = 0.01,
                           eps_irr_sd  = 0.01,
                           start_time = 1990,
                           freq       = 4,
                           nc         = 1,
                           m0         = rep(-3, 2 + 2*nc),
                           level      = 4,
                           slope      = 0.005,
                           breakpoint = NULL,
                           break_level = 0.5,
                           break_drift = FALSE){

  # Generate irreducible error
  error <- ts(rnorm(n, sd = eps_irr_sd), start = start_time, frequency = freq)

  # Generate trend
  eps_zeta <- rnorm(n, sd = eps_zeta_sd)
  # Drift's state
  beta    <- c(slope, rep(0, n))
  for (t in 2:(n + 1)) {

```

```

beta[t] <- ifelse(!is.null(breakpoint)&&(t == breakpoint + 1),
                 break_level^break_drift, 1)*a*beta[t - 1] + eps_zeta[t-1]
# the errors are i.i.d. so t-1 is just for convenient mapping
}

# Level's state
eps_eta <- rnorm(n, sd = eps_eta_sd)

mu <- c(level, rep(0, n))
for (t in 2:(n + 1)) {
  mu[t] <- ifelse(!is.null(breakpoint)&&(t == breakpoint + 1),
                 break_level^!break_drift, 1)*
                 mu[t - 1] + beta[t - 1] + eps_eta[t - 1]
}

# Generate cycle
kappa <- rnorm(n, sd = eps_kappa_sd)
kappa_star <- rnorm(n, sd = eps_kappa_sd)

# low level cycles:
T_matrix <- rho*matrix(c(cos(lambda), sin(lambda),
                        -sin(lambda), cos(lambda)),
                      nrow = 2, byrow = TRUE)
lm <- cbind(rbind(matrix(0, 2, 2*(nc - 1)), diag(2*nc - 2)), matrix(0, 2*nc, 2))
GG <- kronecker(diag(nc), T_matrix) + lm

omegas <- cbind(m0[3:length(m0)], matrix(0, nrow = 2*nc, ncol = n))
kappas <- rbind(kappa, kappa_star, matrix(0, nrow = 2*(nc - 1), ncol = n))

for (t in 2:(n + 1)) {
  omegas[, t] <- GG %*% omegas[, t - 1, drop = F] + kappas[, t - 1, drop = F]
}
omega <- omegas[(nc - 1)*2 + 1,]
omega_star <- omegas[(nc - 1)*2 + 2,]

# High level cycles
C1 <- c(m0[1], rep(0,n))
C2 <- c(m0[2], rep(0,n))

for (t in 2:(n + 1)) {
  C1[t] <- omega[t - 1] + rhoc * C1[t - 1]
  C2[t] <- omega_star[t - 1] + rhoc * C2[t - 1]
}

cycle <- ts(alfa * C1[-1], start = start_time, frequency = freq)
trend <- ts(mu[-1], start = start_time, frequency = freq)
series <- trend + cycle + error

# Return the data
data <- ts.union(series, trend, cycle, error)
options(warn = 0)
invisible(data)
return(data)

```

```

}
# -----

# sim.B_i.gaussian.mixture-----
sim.B_i.gaussian.mixture <- function(N, B_vid, B, plot = FALSE){

  m <- 2
  sigma_1 <- .7
  sigma_2 <- 1.2

  param <- list(c(-1.59, -3.06), c(-1.32, -2.79),
               c(-0.93, -2.36), c(-0.54, -2.02))

  mu_1 <- param[[which(B_vid == B)]] [1]
  mu_2 <- param[[which(B_vid == B)]] [2]

  Z <- sample(c(seq(1,m,1)), size = N, replace = T, prob = c(0.5, 0.5))
  Y_1 <- rnorm(N, mu_1, sigma_1)
  Y_2 <- rnorm(N, mu_2, sigma_2)
  X <- ifelse(Z == 1, Y_1, Y_2)

  if(plot){
    data <- as.data.frame(cbind(Y_1,Y_2,X))
    data <- tidyr::pivot_longer(data, cols = c("Y_1", "Y_2", "X"))

    ggplot(data, mapping = aes(x = value, fill = name)) +
      geom_density(alpha = .25) +
      theme_minimal() +
      scale_fill_discrete(name = "",
                          breaks = c("Y_1", "Y_2", "X"),
                          labels = c("Gaussian 1", "Gaussian 2", "Gaussian Mixture")) +
      xlab("Value") +
      ylab("Density")
  }

  return(X)
}
# -----

# make.tex.tables-----
make.tex.tables <- function(estimates, K, B_vid, rho, N, file_name, not_eval = FALSE){
  x <- matrix(estimates, byrow = TRUE, ncol = length(K)*length(B_vid)*length(rho))
  x_mean <- apply(x, 2, mean, na.rm = TRUE)
  x_sd <- apply(x, 2, sd, na.rm = TRUE)

  mean_tbl <- matrix(round(x_mean,3), ncol = length(rho), byrow = TRUE)

  margin <- qt(0.975,df=N-1)*x_sd/sqrt(N)
  CI_lower<- x_mean - margin
  CI_upper <- x_mean + margin
  CI <- paste("(", round(CI_lower,3), ";", round(CI_upper,3), ")", sep = "")

  CI_tbl <- matrix(CI, ncol = length(rho), byrow = TRUE)

```



```

if(not_eval){
  not_evaluated <- apply(x, 2, function(x){sum(is.na(x))/length(x)})
  not_evaluated_tbl <- matrix(not_evaluated, ncol = length(rho), byrow = TRUE)
  stargazer(not_evaluated_tbl, out=paste0(file_name, "_not_evaluated_proc.tex"))
}

stargazer(mean_tbl, out=paste0(file_name, "_mean.tex"))
stargazer(CI_tbl, out=paste0(file_name, "_CI.tex"))
}
# -----

# make.tex.tables.alpha-----
make_tex_tables_alpha <- function(estimates, K, B_vid, rho, alpha, N,
                                file_name, not_eval = FALSE){
  x <- matrix(unlist(estimates), byrow = TRUE,
             ncol = length(K)*length(B_vid)*length(rho)*length(alpha))
  x_mean <- apply(x, 2, mean, na.rm = TRUE)
  x_sd <- apply(x, 2, sd, na.rm = TRUE)

  mean_tbl <- matrix(round(x_mean,3), ncol = length(rho)*length(alpha), byrow = TRUE)

  margin <- qt(0.975,df=N-1)*x_sd/sqrt(N)
  CI_lower<- x_mean - margin
  CI_upper <- x_mean + margin
  CI <- paste("(", round(CI_lower,3), ":", round(CI_upper,3), ")", sep = "")

  CI_tbl <- matrix(CI, ncol = length(rho)*length(alpha), byrow = TRUE)

  if(not_eval){
    not_evaluated <- apply(x, 2, function(x){sum(is.na(x))/length(x)})
    not_evaluated_tbl <- matrix(not_evaluated, byrow = TRUE,
                              ncol = length(rho)*length(alpha))
    stargazer(not_evaluated_tbl, out=paste0(file_name, "_not_evaluated_proc.tex"))
  }

  stargazer(mean_tbl, out=paste0(file_name, "_mean.tex"))
  stargazer(CI_tbl, out=paste0(file_name, "_CI.tex"))
}
# -----

# est.TTC_PD.params-----
est.TTC_PD.params <- function(country = "DE", scenario = 1, use_hybrid_method = FALSE){

  # read the generated data
  Z_generated <- readRDS(paste0("../04 Generated data/cycle_data_", country, ".RDS"))
  Z_generated <- Z_generated[1:5,]

  # parameters
  N <- nrow(Z_generated)
  K_max <- 5000
  K <- c(100, 500, 1000, 5000)
  TTC_PD <- c(0.01, 0.02, 0.05, 0.1)
  B_vid <- qnorm(TTC_PD)

```

```

rho <- c(0.02, seq(0.05, 0.25, 0.05))
alpha <- 0.5

# matrixes for parameters estimates
rho_MM <- numeric()
B_MM <- numeric()
alpha_MM <- numeric()
rho_reg <- numeric()
B_reg <- numeric()
alpha_reg <- numeric()
ODF_avg <- numeric()
rho_avg_est <- numeric()
B_avg_est <- numeric()
alpha_avg_est <- numeric()
eps_matrix <- matrix(nrow = K_max, ncol = 100)

# #####
for(j in 1:N){

  # common component
  Z <- Z_generated[j, ]

  # idiosyncratic component
  for(k in 1:K_max){
    eps_matrix[k, ] <- rnorm(n=100)
  }

  for(k in 1:length(K)){

    # select the needed number of obligors
    if(K[k] != K_max){
      selected_indexes <- sample(c(1:K_max), K[k], replace = FALSE)
      eps_sub <- eps_matrix[selected_indexes, ]
    } else eps_sub <- eps_matrix

    for(b in 1:length(B_vid)){

      if(scenario == 1){
        B_i <- rnorm(K[k], mean = B_vid[b], sd = 1)
      }

      if(scenario == 2){
        # B_i as mixture of densities
        B_i <- sim.B_i.gaussian.mixture(N = K[k], B_vid = B_vid, B = B_vid[b])
      }

      for(r in 1:length(rho)){

        defaults_matrix <- matrix(nrow = K[k], ncol = 100)
        p_i_hybrid <- matrix(nrow = K[k], ncol = 100)

        for(i in 1:K[k]){

```

```

p_i_hybrid[i, ] <- pnorm((B_i[i] - sqrt(rho[r])*alpha*Z)/
                        sqrt(1-rho[r]*alpha^2))
defaults <- ifelse(sqrt(rho[r])*Z + sqrt(1-rho[r])*eps_matrix[i, ] < B_i[i],
                  1, 0)

defaults_matrix[i, ] <- defaults
}

ODF_t <- apply(defaults_matrix, 2, sum)/K[k]
ODF_t_inv <- qnorm(ODF_t)

# save ODF
ODF_avg <- c(ODF_avg, mean(ODF_t))

# estimation of the parameters rho and B
if((!Inf %in% ODF_t_inv | Inf %in% ODF_t_inv) == FALSE){
  # GMM
  m <- mean(ODF_t_inv)
  sigma <- sd(ODF_t_inv)

  B_vid_hat_mm <- m/sqrt(1+sigma^2)
  rho_hat_mm <- (sigma^2)/(1+sigma^2)

  # TSR
  mdl <- lm(ODF_t_inv ~ 1 + Z)
  B_vid_hat_reg <- mdl$coefficients[[1]]*sqrt(1-rho[r])
  rho_hat_reg <- mdl$coefficients[[2]]^2/(1+mdl$coefficients[[2]]^2)
} else {
  B_vid_hat_mm <- NA
  rho_hat_mm <- NA

  B_vid_hat_reg <- NA
  rho_hat_reg <- NA
}

# save parameter estimates
rho_MM <- c(rho_MM, rho_hat_mm)
rho_reg <- c(rho_reg, rho_hat_reg)

B_MM <- c(B_MM, B_vid_hat_mm)
B_reg <- c(B_reg, B_vid_hat_reg)

# estimation of the parameter alpha
# GMM
p_i_hybrid_inv <- qnorm(p_i_hybrid)
prob_diff <- diff(apply(p_i_hybrid_inv, 2, mean))
Z_diff <- diff(as.vector(Z))

alpha_hat_mm <- mean(sqrt((prob_diff^2)/(rho_hat_mm*(Z_diff^2 + prob_diff^2))))

# TSR

```

```

hybrid_PD_t <- pnorm((B_vid[b] - sqrt(rho[r])*alpha*Z)/sqrt(1-rho[r]*alpha^2))
hybrid_PD_t_inv <- qnorm(hybrid_PD_t)

mdl_2 <- lm(hybrid_PD_t_inv ~ Z)
gamma <- mdl_2$coefficients[[2]]
alpha_hat_reg <- sqrt(gamma^2/((1+gamma^2)*rho_hat_reg))

# save the parameters
alpha_MM <- c(alpha_MM, alpha_hat_mm)
alpha_reg <- c(alpha_reg, alpha_hat_reg)

# hybrid method
if(use_hybrid_method){
  rho_avg <- mean(c(rho_hat_mm, rho_hat_reg))
  B_avg <- mean(c(B_vid_hat_mm, B_vid_hat_reg))

  # TSR
  alpha_hat_reg <- sqrt(gamma^2/((1+gamma^2)*rho_avg))

  # GMM
  alpha_hat_mm <- mean(sqrt((prob_diff^2)/(rho_avg*(Z_diff^2 + prob_diff^2))))

  alpha_avg <- mean(c(alpha_hat_mm, alpha_hat_reg))
}
}
}
}

print(paste("number of iteration:", j, "   time:", lubridate::now()))
}

estimates <- list(rho_GMM = rho_MM,
                 B_GMM = B_MM,
                 alpha_GMM = alpha_MM,
                 rho_TSR = rho_reg,
                 B_TSR = B_reg,
                 alpha_TSR = alpha_reg,
                 ODF = ODF_avg,
                 rho_hybrid = rho_avg_est,
                 B_hybrid = B_avg_est,
                 alpha_hybrid = alpha_avg_est)

return(estimates)
}
# -----

```