

Vogl, Konstantin; Maltritz, Dominik; Huschens, Stefan; Karmann, Alexander

Working Paper

Country Default Probabilities: Assessing and Backtesting

Dresden Discussion Paper Series in Economics, No. 12/06

Provided in Cooperation with:

Technische Universität Dresden, Faculty of Business and Economics

Suggested Citation: Vogl, Konstantin; Maltritz, Dominik; Huschens, Stefan; Karmann, Alexander (2006) : Country Default Probabilities: Assessing and Backtesting, Dresden Discussion Paper Series in Economics, No. 12/06, Technische Universität Dresden, Fakultät Wirtschaftswissenschaften, Dresden

This Version is available at:

<https://hdl.handle.net/10419/22739>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

*Dresden Discussion Paper Series
in Economics*



**Country Default Probabilities:
Assessing and Backtesting**

ALEXANDER KARMANN

STEFAN HUSCHENS

DOMINIK MALTRITZ

KONSTANTIN VOGL

Dresden Discussion Paper in Economics No. 12/06

ISSN 0945-4829

Address of the author(s):

Alexander Karmann
TU Dresden
Faculty of Business Management and Economics
D-01062 Dresden
Germany

e-mail : gkw@mailbox.tu-dresden.de

Stefan Huschens
TU Dresden
Faculty of Business Management and Economics
D-01062 Dresden
Germany

e-mail : Stefan.Huschens@tu-dresden.de

Dominik Maltritz
TU Dresden
Faculty of Business Management and Economics
D-01062 Dresden
Germany

e-mail : Dominik.Maltritz@mailbox.tu-dresden.de

Konstantin Vogl
TU Dresden
Faculty of Business Management and Economics
D-01062 Dresden
Germany

e-mail : Konstantin.Vogl@tu-dresden.de

Editors:

Faculty of Business Management and Economics, Department of Economics

Internet:

An electronic version of the paper may be downloaded from the homepage:
<http://rcswww.urz.tu-dresden.de/wpeconomics/index.htm>

English papers are also available from the SSRN website:
<http://www.ssrn.com>

Working paper coordinator:

Dominik Maltritz
e-mail: wpeconomics@mailbox.tu-dresden.de

Country Default Probabilities: Assessing and Backtesting

Alexander Karmann

TU Dresden

Faculty of Business Management and Economics

D-01062 Dresden

gkw@mailbox.tu-dresden.de

Stefan Huschens

TU Dresden

Faculty of Business Management and Economics

D-01062 Dresden

Stefan.Huschens@tu-dresden.de

Dominik Maltritz

TU Dresden

Faculty of Business Management and Economics

D-01062 Dresden

Dominik.Maltritz@mailbox.tu-dresden.de

Konstantin Vogl

TU Dresden

Faculty of Business Management and Economics

D-01062 Dresden

Konstantin.Vogl@tu-dresden.de

Abstract:

We address the problem how to estimate default probabilities for sovereign countries based on market data of traded debt. A structural Merton-type model is applied to a sample of emerging market and transition countries. In this context, only few and heterogeneous default probabilities are derived, which is problematic for backtesting. To deal with this problem, we construct likelihood ratio test statistics and quick backtesting procedures.

JEL-Classification: C12; C53; F34; G33

Keywords: Sovereign default; Country risk; Default probability; Likelihood ratio test;

1 Introduction

The analysis of country default risk faces several problems. On the one hand, there is the problem of proper estimation of default probabilities. On the other hand, there is the question how to evaluate the estimated probabilities. In this paper, both problems are discussed and solutions are presented.

The difficulties in the estimation process arise from specific features of country risk. The theoretical literature dealing with the causes of defaults of developing countries tells us that – beside the ability to pay – the willingness to pay plays an important role. Both, the ability to pay and – even more – the willingness to pay of countries are rather difficult to quantify based on their causes. Hence, we look on observable effects in terms of prices of defaultable government bonds. These prices are influenced by country default risk and, thus, reflect the market assessment of country risk for the countries considered.

Based on observed market prices we derive the default probability by using a structural credit risk model. Thereby, the parameters of the model have to be estimated simultaneously. Concerning the estimation of the parameters of structural credit risk models for corporate liabilities, there exists a number of papers applying the *two-equations approach*¹ originating from Jones, Mason & Rosenfeld (1984) and Ronn & Verma (1986). This approach can still be found recited in textbooks such as Hull (2003) and Bluhm, Overbeck & Wagner (2003) despite well known objections.² To circumvent this, Duan, Gauthier & Simonato (2005) and Ericsson & Reneby (2005) use a maximum likelihood method valuing corporate liabilities based on the market value of equity.³ This approach is not directly applicable to our case of sovereign risk, as there is no ‘sovereign equity’ at hand. Hence, we estimate the parameters by applying the (modified) maximum likelihood approach to prices of government bonds.

In the light of Basel II, the problem how to backtest default probability forecasts has attracted more and more attention of the banking industry as well as of supervisory authorities. The second part of this paper examines the question to what extent backtesting of heterogeneous default probabilities can be done. A goodness-of-fit test based on the likelihood ratio is constructed for both, testing the two-sided hypothesis that the probability forecasts are exact and testing the one-sided hypotheses that the probability forecasts are not too low or not too high. To determine the quantile of the exact distribution of the likelihood ratio test statistic is a cumbersome task even for small samples. However, for the decision whether to accept or reject the null hypothesis, simply the *p*-value is needed which can be calculated by means of Monte Carlo simulation. Finally, the goodness-of-fit tests developed below are applied to a sample of 105 default probability forecasts for emerging market and transition countries.

The paper is organized as follows: a structural model for assessing country default probabilities is presented in Section 2. The estimation of the model parameters based on market data is described in detail in Section 3. The application and the input data are explained in Section 4. In Section 5, the backtesting problem is addressed. Section 6 concludes.

¹Solving Crosbie & Bohn (2003, p. 16, equation (7) and (8)) or Crouhy, Galai & Mark (2000, p. 88, equation (10) and (11)) for the unknown volatility.

²See Crouhy et al. (2000, p. 88, second paragraph) or Duan (1994, p. 163, footnote 4) for the problematic approach to treat a stochastic volatility as constant.

³This idea is based on Duan (1994), where the underlying framework is formulated rather generally. Duan (1994) provides applications, e.g. for the valuation of insurance contracts for bank deposits, see also the correction Duan (2000).

2 Model

For the following two sections, our focus is on modeling the default triggering process of a single country. The approach is finally applied to 19 different sovereign states. Structural pricing models are based on the idea that a default occurs if a stochastic state variable falls below a certain threshold value. In models considering corporate default risk, the state variable equals the value of the firm.⁴ But what is the appropriate interpretation in the case of sovereign default risk? In the literature, structural pricing models are often applied by estimating the state variable based on possible causes of defaults. Macroeconomic variables are used in order to describe a country's ability to pay. But even if the ability to pay could be measured in a proper way, there is still the problem of unwillingness to pay, because in the case of a sovereign borrower it is not possible to enforce payments internationally. In the present work, market prices are used to infer the value of the stochastic state variable which triggers the default, so that there is no need for an exact distinction between ability and willingness to pay. It can be assumed that the state variable is affected by both.

In the following, the term *default trigger* is used for the stochastic state variable. It is a latent variable which describes the maximum amount of debt servicing payments that a country is able or willing to pay.

To model the stochastic nature of the state variable, the following assumption is made.

Assumption 2.1. *The default trigger is a stochastic process $(W_t)_{t \geq 0}$ which satisfies*

$$dW_t = \mu W_t dt + \sigma W_t dU_t,$$

where μ and $\sigma > 0$ are constants and $(U_t)_{t \geq 0}$ is a standard Brownian motion.

If it is assumed that the debt of the country consists of only one homogeneous class of traded bonds, which promise payments at only one calendar date T of a cumulate amount of B , then the event of default takes a simple form:

Assumption 2.2. *A default occurs if at payment date T the value W_T of the default triggering process is lower than the value B of outstanding debt. In this case, the government pays the amount W_T which it is able and willing to pay. Otherwise, the full amount B is paid out.*

Let $\Phi(\cdot)$ denote the cumulative standard normal distribution function, then the probability of default with time horizon $T - t$ is given by

$$\pi(t, T, \mu, \sigma, B) = \Phi\left(\frac{\ln(B/W_t) - (\mu - \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}}\right). \quad (1)$$

It depends on the actual value of the state variable W_t , its parameters μ and σ , the time to maturity $T - t$ and the contractual payment amount B . Thereby, B and $T - t$ are assumed to be known and exogenously given.

The central part of a structural model is a pricing equation which links the value of default risky securities to its determinants. In structural pricing models the following assumptions are standard:

⁴In the model of Merton (1974), for example, a firm defaults if the value of the firm is below the repayment requirements due.

Assumption 2.3. *Securities are traded without arbitrage opportunities on perfect markets. There is one risk-free interest rate r for borrowing and lending which is constant over time.*

Although idealized, these assumptions are well accepted and used in many models, for example in option pricing theory, real option theory and especially in structural models on corporate as well as sovereign default risk. For the last two applications, the main difference relates to the securities which are considered. In the case of corporate default risk, as in Merton (1974), these securities are the debt contract (derivative security), a risk-free security and the value of the firm (underlying security). In our case, the underlying is the default trigger, which is of course not tradable (and not observable).⁵ But instead of the default trigger, we can use other securities depending on this underlying to build and maintain a risk-free portfolio. For example, credit default swaps or other credit derivatives can be combined with debt contracts to build such a risk-free portfolio. In the last decade, liquid markets for credit default swaps for government bonds have been developed, so that risk-free portfolios can be constructed.

In the framework described above, the fair value at time t of government debt is given by

$$Y_t = Be^{-r(T-t)} - \text{Put}(W_t, \sigma, B, T-t, r),$$

where $\text{Put}(W_t, \sigma, B, T-t, r)$ denotes the Black-Scholes formula for the European put option with underlying W_t , volatility σ , strike B , time to maturity $T-t$ and risk-free interest rate r . Therefore,

$$Y_t = W_t - W_t \Phi(d_t + \sigma\sqrt{T-t}) + Be^{-r(T-t)} \Phi(d_t) \quad (2)$$

where

$$d_t = \frac{\ln(W_t/B) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}.$$

An alternative pricing equation (and a corresponding equation for the default probability) is derived by Geske (1977) for corporate bonds with more than one payment outstanding by using compound option theory.⁶ This makes it possible to model the term structure of debt servicing payments in detail, whereby a default can occur at every payment date. This kind of modeling would be more realistic, since in reality, there is much more than one single payment date. However, the availability of input data is a problem. For example the single payment dates and the corresponding amounts of payment are not known. The only available (actual) data concerning the term structure of debt of developing countries are the amount of short term debt with maturity up to one year and the whole amount of debt (whereas the maturity is not known).

Other papers rely on the approach of Black & Cox (1976), who allow default to occur at any point of time until maturity. This can be seen as a simple model of multi-payment debt, e.g. a coupon bond, where a default is possible on several payment dates. Defaults before maturity are modeled by the assumption that default occurs if the stochastic state variable falls below a certain threshold value for the first time. However, unlike the models of Geske or Merton, this threshold value is not directly linked to the term structure of payments. By contrast, the default threshold is assumed to be a constant or a monotonically increasing function of time. But the debt structure of many countries contradicts this assumption. Typically, for developing countries, the term structure of

⁵It is worth mentioning that the value of a firm is not tradable and usually not observable, too.

⁶See also Geske (1979), Delianedis & Geske (2003) and Maltritz (2006).

debt shows that the short term debt (with a maturity up to one year) is higher than debt repayments due in subsequent years.⁷ Thus, a constant or monotonic threshold value until maturity seems to be rather inadequate. Another drawback of such models is that they involve typically more parameters to be estimated, as for example the value of the default threshold and the recovery rate (respectively the loss given default). Typically, the estimation of these parameters leads to considerable problems.⁸

Hence, like in a number of other papers,⁹ the single payment model explained above is used where only short term debt is considered. Thereby, it is assumed that payments on later payment dates have no influence on the decision or the ability to pay on the first payment date of the short term debt being due. Then, the default threshold equals the amount of short term debt. This kind of application reflects the influence of short term debt on default risk in an appropriate way.

For the model proposed in the present work, the default threshold equals the contractual amount of payments for which data are available. Thus, the debt servicing payments and the repayment date are assumed to be known and exogenously given. Hence, in order to calculate the default probability there are just the values of the default trigger W_t and its parameters μ and σ that need to be determined. How these quantities are estimated is explained in the next section.

3 Parameter Estimation

For each country, we fix time points $t_0 < t_1 < \dots < t_N$ where the market price Y_{t_n} of the country's debt is observed and denoted by Y_n for notational ease. Each risky sovereign bond Y_n is interpreted as a zero coupon bond with nominal value B_n at maturity time T_n . Also, for the latent default trigger, we write W_n instead of W_{t_n} .

The function g_n is defined by

$$g_n(W_n) = B_n e^{-r(T_n - t_n)} - \text{Put}(W_n, \sigma, B_n, T_n - t_n, r), \quad n = 0, \dots, N, \quad (3)$$

so that g_n transforms the underlying W_n into the price Y_n of the risky sovereign bond

$$Y_n = g_n(W_n), \quad n = 0, \dots, N.$$

For fixed quantities σ , B_n , $T_n - t_n$ and r , the Black-Scholes put price is a strictly decreasing function of the underlying, so that the inverse of the function g_n exists, denoted by $h_n = g_n^{-1}$, i.e.

$$W_n = h_n(Y_n), \quad n = 0, \dots, N.$$

Note that the functions g_n and h_n will be considered for different choices of σ .

⁷For the sample of countries analyzed here, the ratio of the short term debt to total debt is 33% on average. If the allocation of residual long term debt on single future years is considered, it can be seen that the actual payment amount becoming due in any single period is much lower.

⁸For example, Keswani (2000) identifies the recovery ratio with the collateralization of Brady bonds. But the payments to bond holders in the case of default can be higher than the collateralization. Lehrbass (2000) assumes a recovery rate of 40%, which is the average recovery rate of US corporate defaults according to *Moody's*, (see Keenan, Carty & Shtogrin (1998)). By contrast, for emerging market countries in the years from 1987 to 2000, the ratio between debt forgiven and debt rescheduled is around 25%, as can be seen from *Global Development Finance*, formerly *World Debt Tables*, published annually by the World Bank, Washington D.C. This means the expected recovery rate should be around 75%. But it is worth mentioning that this ratio is highly volatile. Hence, simple backward looking estimation is problematic.

⁹See, for example, Clark (1991) or Claessens & van Wijnbergen (1993).

Due to the fact that $(W_t)_{t \geq 0}$ is assumed to be a geometric Brownian motion, the log-returns Z_1, \dots, Z_N defined by

$$Z_n = \ln \frac{W_n}{W_{n-1}} = \ln h_n(Y_n) - \ln h_{n-1}(Y_{n-1}), \quad n = 1, \dots, N, \quad (4)$$

are stochastically independent and Z_n has a normal distribution with expectation value $(\mu - \frac{\sigma^2}{2})\tau_n$ and variance $\sigma^2\tau_n$ where

$$\tau_n = t_n - t_{n-1}, \quad n = 1, \dots, N.$$

As shown in Appendix A, the (conditional)¹⁰ log-likelihood function for observed realizations y_0, \dots, y_N of the debt prices Y_0, \dots, Y_N is given by

$$\text{LLF}(\mu, \sigma) = - \sum_{n=1}^N \left[\ln \sqrt{2\pi\sigma^2\tau_n} + k_n(y_n) + \frac{\left(\ln \frac{h_n(y_n)}{h_{n-1}(y_{n-1})} - (\mu - \frac{\sigma^2}{2})\tau_n \right)^2}{2\sigma^2\tau_n} \right]$$

where the term

$$k_n(y_n) = \ln h_n(y_n) + \ln \Phi \left(\frac{\ln \frac{B_n}{h_n(y_n)} - (r + \frac{\sigma^2}{2})(T_n - t_n)}{\sigma \sqrt{T_n - t_n}} \right)$$

reflects that the log-returns Z_n are unobservable and that the actual observations are the transformed data Y_n .

Since the functions h_n and k_n depend on σ but not on μ , the log-likelihood function is maximized for a fixed σ with respect to μ by

$$\mu^*(\sigma) = \frac{\sigma^2}{2} + \frac{\sum_{n=1}^N \ln \frac{h_n(y_n)}{h_{n-1}(y_{n-1})}}{\sum_{n=1}^N \tau_n}. \quad (5)$$

Note that $\sum_{n=1}^N \tau_n = t_N - t_0$ and that the value of $h_n(y_n)$ can be determined numerically for every fixed σ . Numerical maximization of the function $\text{LLF}(\mu^*(\sigma), \sigma)$ with respect to σ yields the maximum likelihood estimate

$$\hat{\sigma}_{\text{ML}} = \underset{\sigma}{\operatorname{argmax}} \text{LLF}(\mu^*(\sigma), \sigma)$$

for σ . The maximum likelihood estimate for μ is then given by

$$\hat{\mu}_{\text{ML}} = \mu^*(\hat{\sigma}_{\text{ML}}).$$

Given these estimates of μ and σ , the actual time t and the observed value of the current market price Y_t of the country's debt, equation (2) can be solved numerically for the trigger W_t . Then, the probability that the country will be in default at maturity time $T - t$ can be calculated by (1).

We add two remarks regarding the existing literature. First, the maximum likelihood estimation procedure can be applied to infer the drift μ and the volatility σ of the latent underlying process from price observations of any contingent claim.¹¹ In the present work, bond prices are used for the parameter estimation, whereas in Duan et al. (2005)

¹⁰On the condition that $Y_0 = y_0$.

¹¹See Duan (1994).

the parameters of the corresponding model are derived from stock prices which, in the Merton model, are interpreted as call prices. No matter whether the Black-Scholes bond prices or stock prices are used, in both approaches, the log-likelihood can analytically be maximized with respect to μ as in (5), so that numerical maximization has to be performed in one dimension only, i.e. with respect to σ .¹² Second, in both approaches the maximum likelihood estimate for σ could also be obtained by using the following iterative procedure: choose an initial value for σ , compute the implied time series of log-returns z_1, \dots, z_N as in (4), update σ by the sample standard deviation of z_1, \dots, z_N , compute the implied time series of log-returns again, and so on. This procedure is heuristically mentioned in Crosbie & Bohn (2003) from *Moody's KMV*. But it was shown by Duan et al. (2005) that in the basic Merton model this iteration will also converge to the maximum likelihood estimate of σ , and that the iterative procedure is equivalent to determining the ML estimate by an EM algorithm. In more complex models however, the application of the iterative procedure mentioned above has no theoretical justification and may indeed lead to poor estimates.

4 Data and Application

Our estimation approach is now applied to a sample of 19 emerging market and transition countries to estimate default probabilities for quarterly data from 1998 till 2005. The observation dates t_0, t_1, \dots, t_N are identified with the first trading date in each quarter. Thus, B_n is the amount of debt with a maturity up to one year (i.e. the following four quarters) outstanding at the beginning of the n th quarter, i.e. $T_n = t_{n+4}$. How the corresponding market value Y_n of this debt at time t_n and the other input data are determined is explained now.

The following data have to be known for each country and each estimation date: market data on prices respectively interest rates of government bonds, data on the amount of (short term) repayment requirements and, finally, the risk-free interest rate. The selection of countries and observation periods are in large part determined by the availability of these input data.

Data on sovereign debt strongly limit the frequency of the time series used as well as the modeling of the term structure of debt. As mentioned earlier, current and detailed information on the duration of single debt issues or on the amounts of debt corresponding to certain duration periods are not available. Only for short term debt (up to one year) the amount due is published. For the long term debt, no detailed and current data are available. Hence, we focus on short term debt and short term default probability (as explained in Section 2) and assume the time to maturity to be one year.

Current data on country debt are provided by different sources. For example, there are annual data from the *Economist Intelligence Unit*[®] (EIU), which are available via *DataStream*[®], and quarterly data from the joint database of international organizations (BIS, IMF, OECD, World Bank). The main drawback of the joint database is that no data on the whole amount of short term debt are available. For example, trade credits or credits to other countries are not provided. Nevertheless, this database enables the market participants to obtain information on the changes of debt, so that rational agents are able to infer correct market prices of risk on a quarterly base. According to the availability of data, our methodology is applied quarterly, whereas total short term debt is updated on an annual base, as provided by the EIU.

¹²This point is not clarified by Duan (1994, p.165).

The amount of debt B_T and the time to maturity $T - t$ are needed in both, the equation for the default probability (1) and the pricing equation (2). The latter is used to determine the other quantities, i.e. the default triggering process and its parameters. To estimate these quantities, the market value of the debt, Y_t , is needed. The market value of debt is derived from data on the market prices of government bonds issued by the respective states. More precisely, we use the internal rates of return as given by the market prices of the bonds.

Generally, the current value of future payments is calculated by discounting these promised future payments. If there is no default risk, the risk-free interest rate should be applied. Since the payments considered here are affected by default risk, a risk-adjusted interest rate is to be used. This risk-adjusted interest rate is derived from the internal rate of return of market traded government bonds.

Here we use data from the *Emerging Market Bond Index*⁺ (EMBI⁺) published by *JP Morgan*. The EMBI⁺ is calculated for a number of countries based on Brady- and Eurobonds and represents the weighted average of several single bond issues (in case more than one issue of a country exists). To calculate the market value Y_t of the country debt considered at date t , the (short term) repayment requirements B are discounted over the time to maturity $T - t$ of one year using the risky interest rate given by the EMBI⁺.

Besides the amount of debt and its market value, the risk-free interest rate is needed in the pricing equation (2). It is derived by subtracting the spread from the corresponding risky interest rate. This procedure is based on Assumption 2.3 which implies a flat term structure of interest rates.

Data on the spread are also provided by *JP Morgan* as part of the EMBI⁺ data. According to *JP Morgan*, the published spread is the difference between the average risky interest rate and the interest rate given by the US term structure which corresponds to the average time to maturity of the bonds used to calculate the EMBI⁺. Thereby, it is assumed that the default risk of US bonds is negligible and, hence, the corresponding interest rates are unaffected by default risk.

For a given time series of market values of debt Y_n , ($n = 0, \dots, N$), as well as the corresponding risk-free interest rates r_n and debt values B_n , the estimation methodology described in the last section is used to derive the estimates of the model parameters μ and σ , and of the default trigger W_n . Hence, all quantities needed to calculate the default probability via formula (1) are available.

The analyzed sample of countries as well as of the observation period is determined by the availability of data from EMBI⁺. To achieve time series of sufficient length we consider 19 countries for which EMBI⁺ data are available since 1998. For countries where the necessary data are available at earlier dates the estimation includes these observations, too.

Table 1 shows the annual default probabilities forecasted at the beginning of every year. For the evaluation of these probabilities, we need a definition of crisis. Our notion of debt crisis, respectively default, follows Manasse, Roubini & Schimmelpfennig (2003). A country is defined to be in default if it has a rating of D or SD by *Standard & Poor's* or if a large IMF emergency credit (which is below 100 percent of quota) is announced. The latter is applied in order to cover situations where a country is de facto in default while a default in the legal sense is avoided by the help of the international community. We assume a debt crisis to persist as long as the *S&P* rating is D or SD, or as long as the IMF arrangement lasts or is renewed. Years where the country already is in default at the first of January are excluded since it is the aim of the structural approach to estimate the probability of default and not the probability of remaining in a debt crisis.

In addition, we add a waiting period of one year before we continue the calculation and evaluation of default probabilities. This accounts for the fact that a default in the year following a debt crisis is seen as a continuation of the existing crisis rather than as an arising of a new crisis. Periods which are not considered in the evaluation are symbolized by missing values (dots) in Table 1. The numbers in Table 1 describe the probability that the observed country will default during the year considered. These annual default probabilities are used for the evaluation performed in the next section.

Table 1: Estimated annual default probabilities

| Country | Year | | | | | | | |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 |
| Argentina | 0.112 | 0.162 | 0.126 | 0.177 | . | . | . | . |
| Brazil | 0.162 | . | . | . | . | . | . | . |
| Bulgaria | . | . | . | . | . | 0.059 | 0.036 | 0.018 |
| China | 0.012 | 0.020 | 0.013 | 0.016 | 0.009 | 0.008 | 0.005 | 0.005 |
| Colombia | 0.212 | 0.357 | . | . | . | . | . | . |
| Ecuador | 0.065 | 0.159 | . | . | 0.122 | 0.173 | 0.080 | 0.068 |
| Malaysia | 0.037 | 0.108 | 0.032 | 0.044 | 0.038 | 0.038 | 0.018 | 0.015 |
| Mexico | 0.145 | 0.233 | 0.136 | 0.139 | 0.113 | 0.117 | 0.075 | 0.066 |
| Morocco | 0.120 | 0.189 | 0.109 | 0.155 | 0.137 | 0.101 | 0.043 | 0.054 |
| Panama | . | 0.135 | 0.120 | 0.144 | 0.121 | 0.128 | 0.101 | 0.098 |
| Peru | . | . | 0.143 | 0.207 | 0.165 | 0.183 | 0.105 | 0.090 |
| Philippines | . | . | . | . | 0.230 | 0.236 | 0.200 | 0.220 |
| Poland | 0.011 | 0.013 | 0.009 | 0.012 | 0.009 | 0.008 | 0.003 | 0.002 |
| Russia | 0.065 | . | . | . | 0.065 | 0.044 | 0.022 | 0.020 |
| South Africa | 0.048 | 0.124 | 0.053 | 0.081 | 0.062 | 0.045 | 0.027 | 0.018 |
| South Korea | . | . | . | . | 0.038 | 0.025 | 0.017 | 0.020 |
| Thailand | . | . | . | . | 0.007 | 0.006 | 0.003 | 0.002 |
| Turkey | 0.155 | 0.270 | 0.174 | . | . | . | . | . |
| Venezuela | 0.096 | 0.236 | 0.179 | 0.195 | 0.222 | 0.221 | 0.119 | 0.089 |

5 Backtesting

Problem and data

Predicted default probabilities, as given in Table 1, need some kind of evaluation. Our focus is not on testing binary default predictions, as widely used in the literature¹³, but on evaluation methods for default probabilities. Thereby, the method shall be applicable especially in the case of heterogeneous default probabilities where any attempt to construct nearly homogeneous classes (each with a reasonable number of members) fails.

More precisely, we deal with estimated default probabilities being the components of the vector

$$\hat{\boldsymbol{\pi}} = (\hat{\pi}_1, \hat{\pi}_2, \dots, \hat{\pi}_m).$$

Here, $\hat{\boldsymbol{\pi}}$ consists of the $m = 105$ annual default probability estimates given in Table 1. The true, but unknown, default probabilities

$$\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_m)$$

¹³See Pesaran & Timmermann (1992).

are unobservable. Observable are the outcomes x_i of Bernoulli random variables X_i as given in Table 2 with

$$P(X_i = 1) = \pi_i, \quad P(X_i = 0) = 1 - \pi_i, \quad i = 1, \dots, m,$$

where $X_i = 1$ if the debtor is in default at the end of the year, and $X_i = 0$ if not. The binary default indicator is reported only for those country-year combinations where default probabilities have been estimated in Table 1. The goal is to find a statistical test

Table 2: Observed realizations x_i of Bernoulli variables X_i

| Country | Year | | | | | | | |
|--------------|------|------|------|------|------|------|------|------|
| | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 |
| Argentina | 0 | 0 | 0 | 1 | . | . | . | . |
| Brazil | 1 | . | . | . | . | . | . | . |
| Bulgaria | . | . | . | . | . | 0 | 0 | 0 |
| China | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Colombia | 0 | 1 | . | . | . | . | . | . |
| Ecuador | 0 | 1 | . | . | 0 | 0 | 0 | 0 |
| Malaysia | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mexico | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Morocco | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Panama | . | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Peru | . | . | 0 | 0 | 0 | 0 | 0 | 0 |
| Philippines | . | . | . | . | 0 | 0 | 0 | 0 |
| Poland | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Russia | 1 | . | . | . | 0 | 0 | 0 | 0 |
| South Africa | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| South Korea | . | . | . | . | 0 | 0 | 0 | 0 |
| Thailand | . | . | . | . | 0 | 0 | 0 | 0 |
| Turkey | 0 | 0 | 1 | . | . | . | . | . |
| Venezuela | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

in order to decide whether the outcomes x_1, \dots, x_m of the Bernoulli trials X_1, \dots, X_m are compatible with the stated default probabilities $\hat{\pi}_1, \dots, \hat{\pi}_m$ or whether the hypothesis $\pi = \hat{\pi}$ should be rejected.

Likelihood ratio tests

From a statistical point of view, the numbers $\hat{\pi}_i$ are estimates for the unknown distribution parameters π_i . The difference to the situation usually treated in textbooks on statistical inference is that the X_i are not identically distributed, but each X_i is governed by its own Bernoulli parameter π_i . In the common framework of hypothesis testing relying on asymptotic statistics the observations are usually drawn from the same distribution. Here, in contrast, by observing more and more X_i , we increase the number of distribution parameters to be tested. With each additionally observed X_i , a new unknown Bernoulli parameter enters the hypothesis to be inspected.

Thus, there is no hope to find a test that can detect when one of the $\hat{\pi}_i$ has been a wrong prediction, since, strictly speaking, we have only one single observation X_i for each

distribution. Nonetheless, the prediction system of default probabilities can be tested as a whole. For a first attempt, we assume X_1, \dots, X_m to be stochastically independent.¹⁴

Theorem 5.1. *The test for the hypothesis*

$$\mathcal{H}_0 : \boldsymbol{\pi} = \hat{\boldsymbol{\pi}} \quad \text{versus} \quad \mathcal{H}_1 : \boldsymbol{\pi} \neq \hat{\boldsymbol{\pi}} \quad (6)$$

with test statistic

$$T_1 = \sum_{i=1}^m X_i \ln \frac{\hat{\pi}_i}{1 - \hat{\pi}_i} \quad (7)$$

is a likelihood ratio test of size

$$\alpha_k = P_{\hat{\boldsymbol{\pi}}}(T_1 < k),$$

where \mathcal{H}_0 is rejected if $T_1 < k$.

The subscript $\hat{\boldsymbol{\pi}}$ of P denotes the case where the vector $\boldsymbol{\pi}$ of true unknown default probabilities equals $\hat{\boldsymbol{\pi}}$. Choosing the value of k determines the level of significance α_k . The proof of Theorem 5.1 is given in Appendix B. The exact distribution of T_1 under \mathcal{H}_0 can be calculated in principle. But finding the number k which yields a given significance level α , might become a cumbersome computational task when m is large: there are 2^m points in $\{0, 1\}^m$ which all have positive probability and are relevant for computational inspection. Instead, for computational ease, we calculate the p -value using the Monte Carlo method. The p -value is the smallest possible significance level so that \mathcal{H}_0 is rejected based on the observed values x_1, \dots, x_m . Applying the test to the data shown in Tables 1 and 2, we get the test statistic

$$\sum_{i=1}^{105} x_i \ln \frac{\hat{\pi}_i}{1 - \hat{\pi}_i} = -9.65728.$$

Running 10^8 Monte Carlo trials¹⁵, we obtain the following p -value

$$P_{\hat{\boldsymbol{\pi}}} \left(\sum_{i=1}^{105} X_i \ln \frac{\hat{\pi}_i}{1 - \hat{\pi}_i} \leq -9.65728 \right) = 0.9376 \text{ (0.000024)}$$

with an estimated standard error of $2.4 \cdot 10^{-5}$. The test can be used in order to detect significant deviations from the null hypothesis $\mathcal{H}_0 : \boldsymbol{\pi} = \hat{\boldsymbol{\pi}}$. For our sample, i.e. anyone questioning the probability estimates $\hat{\boldsymbol{\pi}}$ and advocating the alternative \mathcal{H}_1 must admit a probability of type I error of at least 93%.

In practical applications, e.g. taking the position of a banking supervisory authority, it may be uncritical if banks overestimate the default probabilities of their debtors. Here, a one-sided test may be of interest. By rejecting the null hypothesis of the following test, an authority could prove that a bank significantly underestimates some of its default probabilities.

¹⁴See the discussion on this point below.

¹⁵This number of Monte Carlo trials assures that the p -value is determined with an accuracy of $\pm 10^{-4}$. Thereby, the accuracy is measured by four standard deviations. For practical purposes, an accuracy of 10^{-3} or even 10^{-2} may be sufficient, which requires only $n = 10^6$ or $n = 10^4$ Monte Carlo trials, considering that the standard deviation is proportional to $n^{-1/2}$.

Theorem 5.2. *The test for the hypothesis*

$$\mathcal{H}_0 : \boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}} \quad \text{versus} \quad \mathcal{H}_1 : \boldsymbol{\pi} \not\leq \hat{\boldsymbol{\pi}} \quad (8)$$

with test statistic

$$T_2 = \sum_{i=1}^m X_i \ln \hat{\pi}_i \quad (9)$$

is a likelihood ratio test of size

$$\alpha_k = \sup_{\boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}} P_{\boldsymbol{\pi}}(T_2 < k) = P_{\hat{\boldsymbol{\pi}}}(T_2 < k),$$

where \mathcal{H}_0 is rejected if $T_2 < k$.

Thereby, the relation \leq between vectors means that \leq holds for all components. In contrast, $\not\leq$ for vectors means that at least one component violates the relation \leq . Theorem 5.2 is proved in Appendix C.

Testing the hypotheses (8) for the probability estimates of Table 1 with the data in Table 2 yields the test statistic

$$\sum_{i=1}^{105} x_i \ln \hat{\pi}_i = -10.9018.$$

The p -value

$$P_{\hat{\boldsymbol{\pi}}} \left(\sum_{i=1}^{105} X_i \ln \hat{\pi}_i \leq -10.9018 \right) = 0.9347 \text{ (0.000025)}$$

is determined by 10^8 Monte Carlo trials with an estimated standard error of $2.5 \cdot 10^{-5}$. This test can be used to detect significant deviations from the null hypothesis $\mathcal{H}_0 : \boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}$. In our case, anyone advocating \mathcal{H}_1 , i.e. supporting the opinion that some of the probability estimates $\hat{\boldsymbol{\pi}}$ are too low, must admit a probability of type I error of at least 93%.

For a bank as a lending institution, it is not of its own interest to use default probability estimates that are too large, thereby tying more regulatory capital than necessary. Anyone using default probability estimates that are significantly too large is alerted when rejecting the null hypothesis of the following test.

Theorem 5.3. *The test for the hypothesis*

$$\mathcal{H}_0 : \boldsymbol{\pi} \geq \hat{\boldsymbol{\pi}} \quad \text{versus} \quad \mathcal{H}_1 : \boldsymbol{\pi} \not\geq \hat{\boldsymbol{\pi}} \quad (10)$$

with test statistic

$$T_3 = \sum_{i=1}^m (1 - X_i) \ln(1 - \hat{\pi}_i) \quad (11)$$

is a likelihood ratio test of size

$$\alpha_k = \sup_{\boldsymbol{\pi} \geq \hat{\boldsymbol{\pi}}} P_{\boldsymbol{\pi}}(T_3 < k) = P_{\hat{\boldsymbol{\pi}}}(T_3 < k),$$

where \mathcal{H}_0 is rejected if $T_3 < k$.

Thereby, the relations \geq and $\not\geq$ between vectors are analogously defined as above. Theorem 5.3 follows directly from Theorem 5.2 taking into account that $\boldsymbol{\pi} \geq \hat{\boldsymbol{\pi}}$ is equivalent to $1 - \boldsymbol{\pi} \leq 1 - \hat{\boldsymbol{\pi}}$.

Applying this test to the probability estimates of Table 1 and the data in Table 2, we get the test statistic

$$\sum_{i=1}^{105} (1 - x_i) \ln(1 - \hat{\pi}_i) = -9.52120.$$

From 10^8 Monte Carlo trials we infer the p -value

$$P_{\hat{\boldsymbol{\pi}}} \left(\sum_{i=1}^{105} (1 - X_i) \ln(1 - \hat{\pi}_i) \leq -9.52120 \right) = 0.1982 \text{ (0.000040)}$$

with an estimated standard error of $4.0 \cdot 10^{-5}$. Anyone advocating \mathcal{H}_1 , i.e. thinking that some of the probability estimates $\hat{\boldsymbol{\pi}}$ are too large, must admit a probability of type I error of at least 19%.

In practical applications, depending on the respective problem, one would choose one of the three tests. Here, the results of the three tests together indicate overall confirmation of the predicted default probabilities. Indeed, the fact that none of the three null hypotheses can be rejected with any usual error probability suggests that our default probabilities are highly consistent with the data.

Quick tests based on asymptotics

Since the test statistics (7), (9) and (11) are regarded as sums of stochastically independent random variables, the question arises whether the central limit theorem applies, so that the standard normal distribution could be used to infer approximate p -values instead of performing Monte Carlo simulations. Assuming that $\boldsymbol{\pi} = \hat{\boldsymbol{\pi}}$, as postulated by \mathcal{H}_0 in (6), the standardized test statistic is

$$\tilde{T}_1 = \frac{T_1 - \mu_1}{\sigma_1}, \quad (12)$$

where the expectation and variance of T_1 according to (7) are given by

$$\mu_1 = \sum_{i=1}^m \hat{\pi}_i \ln \frac{\hat{\pi}_i}{1 - \hat{\pi}_i} \quad \text{and} \quad \sigma_1^2 = \sum_{i=1}^m \hat{\pi}_i (1 - \hat{\pi}_i) \left(\ln \frac{\hat{\pi}_i}{1 - \hat{\pi}_i} \right)^2.$$

The likelihood ratio test of Theorem 5.1 can equivalently be applied using the standardized test statistic \tilde{T}_1 . Under the condition

$$\lim_{m \rightarrow \infty} \sum_{i=1}^m \hat{\pi}_i (1 - \hat{\pi}_i) = \infty,$$

the sample distribution of \tilde{T}_1 under \mathcal{H}_0 converges to the standard normal distribution.¹⁶ Note that even though the standard normal distribution is symmetric, the null hypothesis in (6) is rejected only for low values of \tilde{T}_1 , i.e. values which are less than the chosen threshold value, since low values of \tilde{T}_1 imply a low likelihood ratio.

The tests of Theorem 5.2 and 5.3 can analogously be based on standardized test statistics \tilde{T}_2 and \tilde{T}_3 . For the probability estimates in Table 1, the sample distributions of \tilde{T}_1 , \tilde{T}_2 and \tilde{T}_3 given $\boldsymbol{\pi} = \hat{\boldsymbol{\pi}}$ are compared with the standard normal approximation in

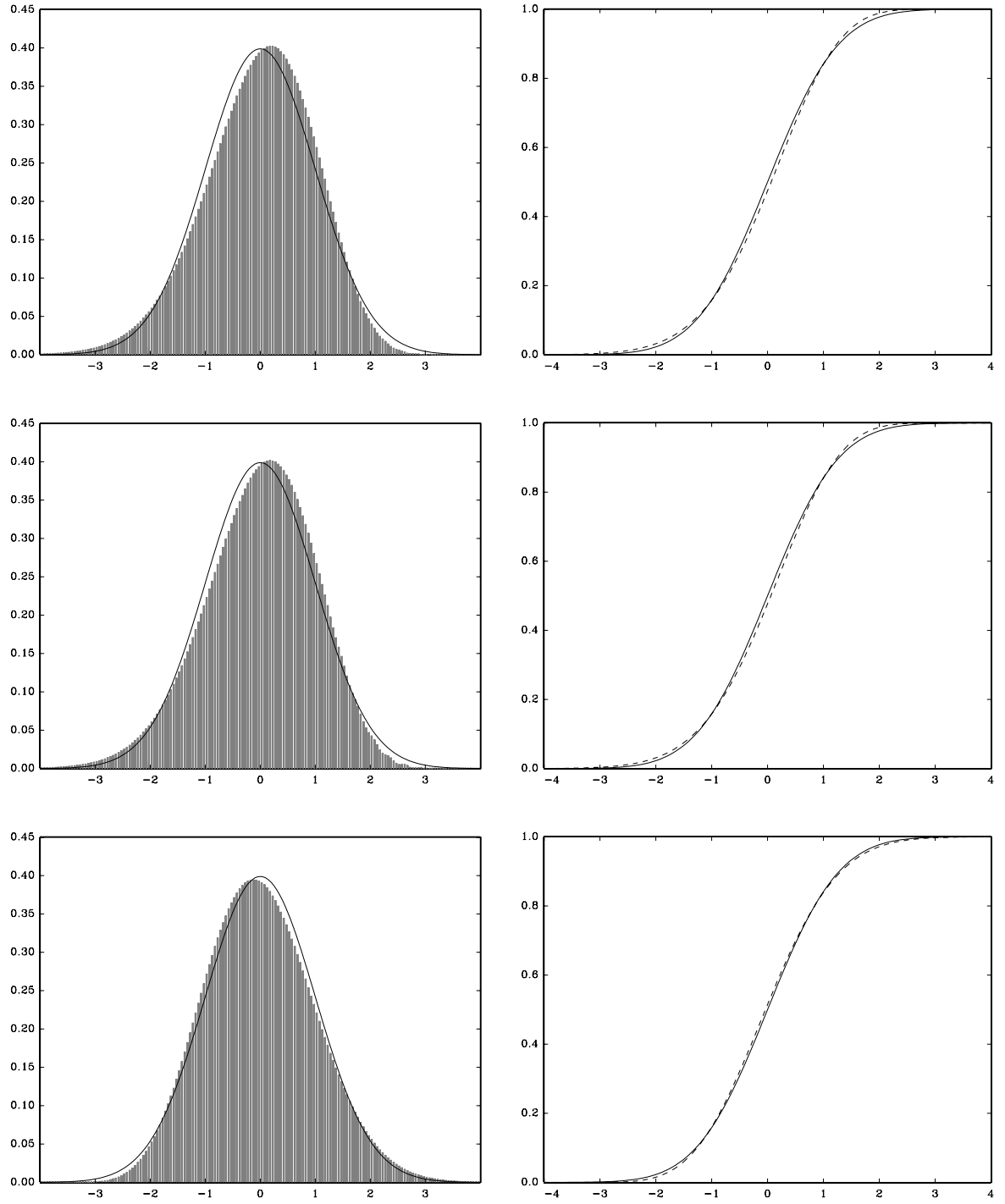


Figure 1: Sample distributions of the test statistics \tilde{T}_1 (top), \tilde{T}_2 (center) and \tilde{T}_3 (bottom) compared to the density (left panels) and the distribution function (right panels) of the standard normal distribution (solid line)

Figure 1. In the left panels the sample distributions from 10^8 Monte Carlo iterations are compared to the standard normal density. In the right panels the respective cumulative distribution functions are compared to the standard normal distribution function.

In Table 3, the p -values of the likelihood ratio tests, applied to the default probability estimates of Table 1 and the observations of Table 2, are reported together with their standard normal approximation. The effect of the central limit theorem seems to

Table 3: Exact p -values and standard normal approximations for the likelihood ratio tests applied to the default probabilities of Table 1 and the data of Table 2

| Testing \mathcal{H}_0 : | p -value from | |
|------------------------------|--------------------------|-------------------------|
| | simulation | standard |
| | of exact distribution | normal approximation |
| $\pi = \hat{\pi}$ | 0.9376 | 0.9260 |
| $\pi \leq \hat{\pi}$ | 0.9347 | 0.9243 |
| $\pi \geq \hat{\pi}$ | 0.1982 | 0.1927 |

be strong enough in order to calculate approximate p -values from the standard normal distribution. We conclude that a quick test based on the standard normal approximation is justified here. In general however, the question whether the normal approximation of the distribution of the likelihood ratio statistic is appropriate, depends on m and the specific values of $\hat{\pi}$.

Methodological Aspects

The assumption that the Bernoulli trials X_i are stochastically independent is critical, since i enumerates all country-year combinations of the estimated annual default probabilities. Whereas cross-sectional independence among the countries within a one-year period is rather an idealization, serial independence over time can be justified by an argument relying on conditional independence, considering the fact that the default probabilities are conditional on the information of the past history.

The methods presented in Section 5 may serve as goodness-of-fit tests applicable without large computational effort. They do not rely on the model in which the default probabilities are estimated, instead they are based on the assumption of independent default events. However, backtesting of default probabilities which accounts for dependencies cannot be formulated in this manner. The modeling of default correlations is usually based on conditional independence given some exogeneous or endogeneous risk factors. Therefore, understanding the case of independent Bernoulli trials is an inevitable requisite for any credit risk model. To detect parameter misspecification in such a correlation model, a lot more observations than in our case are necessary, firstly because additional parameters specifying the dependence structure must be estimated, and secondly because more extreme realizations x_1, \dots, x_m are compatible with the hypothesis $\pi_1 = \hat{\pi}_1, \dots, \pi_m = \hat{\pi}_m$.

Under the assumption of independent Bernoulli trials, it can be shown that no other tests can be found yielding uniformly better error probabilities, this means, the tests are

¹⁶See Petrov (1995, pp. 123–125).

admissible¹⁷ in the sense of statistical test theory. The criterion of unbiasedness¹⁸ for hypothesis testing cannot be applied because, except for trivial and non-relevant cases, no unbiased tests exist for the hypotheses considered here. If the null hypothesis \mathcal{H}_0 is exchanged for its alternative \mathcal{H}_1 in (6),(8) or (10) respectively, no non-trivial tests exist at all. Thus, we cannot confirm the predicted default probabilities by rejecting an alternative hypothesis with a given error probability, but we can – more weakly – confirm the predicted probabilities by not rejecting the null hypothesis which is a typical property of goodness-of-fit tests.

6 Conclusion

This paper deals with several problems arising in the analysis of country default risk. This concerns the estimation of default probabilities as well as their evaluation. According to the theoretical literature, besides the ability to pay the willingness to pay is an important determinant of country default. Since both, the ability to pay as well as the willingness to pay are difficult to estimate based on their – partly unobservable – causes, we quantify the default risk based on observable effects: we use the prices, respectively the corresponding risky interest rates, of defaultable government bonds because they reflect the market assessment of default risk. To infer the default risk from the bond market data we estimate a structural pricing model. This is done for a broad sample of emerging market and transition countries for which the necessary market data are available.

To our knowledge, the problem how to backtest heterogeneous default probabilities has not yet been treated in the literature. The problem is to evaluate not predictions of events but probability predictions. We suggest goodness-of-fit tests based on likelihood ratio test statistics. The tests can be applied even if the sample size is small. To derive the test decision, the p -value can be obtained by Monte Carlo simulation with high accuracy and little computational effort. A standard normal approximation is justified when the sample size is large, even more simplifying the computational effort. The Revised International Capital Framework of Basel II requires the validation of all internal estimates of risk parameters, especially the backtesting of estimated default probabilities.¹⁹ If the number of debtors is small and default probabilities are heterogeneous, which is typical for country risk, the proposed methodology can be applied where standard approaches²⁰ fail.

The evaluation of the estimated default probabilities yields plausible and promising results. The use of structural pricing models and the maximum likelihood approach to estimate the unknown quantities based on market data is a practical methodology to quantify country default risk. The methods presented in the paper give also a solution to the problem how to backtest country default probabilities, more generally speaking, to tackle the evaluation problem when default probabilities are heterogeneous.

¹⁷See e.g. Lehmann & Romano (2005, p. 17).

¹⁸See e.g. Lehmann & Romano (2005, p. 110).

¹⁹See Basel Committee on Banking Supervision (2006, pp. 109–110) and Basel Committee on Banking Supervision (2005, pp. 8–9).

²⁰For example the binomial and chi-square test mentioned in Basel Committee on Banking Supervision (2005, p. 33).

A Appendix. LLF for Transformed Data

For technical reasons, we add the definition

$$Z_0 = \ln W_0 = \ln h_0(Y_0) \quad (13)$$

and assume Z_1, \dots, Z_N to be stochastically independent of Z_0 . The probability density function of Z_0, \dots, Z_N is given by

$$f_{Z_0, \dots, Z_N}(z_0, \dots, z_N) = f_{Z_0}(z_0) \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma^2\tau_n}} \exp \left\{ -\frac{1}{2} \left(\frac{z_n - (\mu - \frac{\sigma^2}{2})\tau_n}{\sigma\sqrt{\tau_n}} \right)^2 \right\}$$

The random vector Y_0, \dots, Y_N is a transformation of Z_0, \dots, Z_N , the backward transformation is given by (13) and (4). Thus, the probability density function of Y_0, \dots, Y_N is given by²¹

$$\begin{aligned} f_{Y_0, \dots, Y_N}(y_0, \dots, y_N) &= \\ &= \left| \det(J) \right| \cdot f_{Z_0}(\ln h_0(y_0)) \cdot \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma^2\tau_n}} \exp \left\{ -\frac{1}{2} \left(\frac{\ln \frac{h_n(y_n)}{h_{n-1}(y_{n-1})} - (\mu - \frac{\sigma^2}{2})\tau_n}{\sigma\sqrt{\tau_n}} \right)^2 \right\} \end{aligned}$$

where J is a Jacobi matrix whose elements above and right to the main diagonal equal zero. The n th element on the main diagonal of J is

$$J_{nn} = \frac{h'_n(y_n)}{h_n(y_n)} = \frac{1}{h_n(y_n)} \frac{1}{g'_n(h_n(y_n))} \quad n = 0, \dots, N,$$

where the prime symbol $'$ stands for the derivative. Recall equation (3) for the fact that the function g'_n equals the “delta” of a European put option multiplied by -1 , so that

$$g'_n(h_n(y_n)) = 1 - \Phi \left(\frac{\ln \frac{h_n(y_n)}{B_n} + (r + \frac{\sigma^2}{2})(T_n - t_n)}{\sigma\sqrt{T_n - t_n}} \right).$$

It is

$$\left| \det(J) \right| = \prod_{n=0}^N J_{nn}$$

by virtue of the zero entries in J and the positivity of the functions h_n and g'_n . Taking into account that

$$f_{Y_0}(y_0) = J_{00} f_{Z_0}(\ln h_0(y_0)),$$

the density of Y_0, \dots, Y_N can be written as

$$\begin{aligned} f_{Y_0, \dots, Y_N}(y_0, \dots, y_N) &= \\ &= f_{Y_0}(y_0) \cdot \prod_{n=1}^N \frac{J_{nn}}{\sqrt{2\pi\sigma^2\tau_n}} \exp \left\{ -\frac{1}{2} \left(\frac{\ln \frac{h_n(y_n)}{h_{n-1}(y_{n-1})} - (\mu - \frac{\sigma^2}{2})\tau_n}{\sigma\sqrt{\tau_n}} \right)^2 \right\}. \end{aligned}$$

²¹See e.g. Casella & Berger (2002, p. 185).

Hence, the conditional²² log-likelihood function for observations y_0, \dots, y_N of debt prices is given by

$$\begin{aligned} \text{LLF}(\mu, \sigma) &= \ln f_{Y_1, \dots, Y_N | Y_0}(y_1, \dots, y_N | y_0) = \\ &= - \sum_{n=1}^N \left[\ln \frac{\sqrt{2\pi\sigma^2\tau_n}}{J_{nn}} + \frac{\left(\ln \frac{h_n(y_n)}{h_{n-1}(y_{n-1})} - (\mu - \frac{\sigma^2}{2})\tau_n \right)^2}{2\sigma^2\tau_n} \right] \\ &= - \sum_{n=1}^N \left[\ln \sqrt{2\pi\sigma^2\tau_n} + \ln \left(h_n(y_n) g'_n(h_n(y_n)) \right) + \frac{\left(\ln \frac{h_n(y_n)}{h_{n-1}(y_{n-1})} - (\mu - \frac{\sigma^2}{2})\tau_n \right)^2}{2\sigma^2\tau_n} \right]. \end{aligned}$$

B Appendix. Proof of Theorem 5.1

Let $\mathbf{x} = (x_1, \dots, x_m) \in \{0, 1\}^m$ be a fixed vector of outcomes of independent Bernoulli random variables X_1, \dots, X_m . The likelihood function is

$$\mathbf{L}(\boldsymbol{\pi}; \mathbf{x}) = \prod_{i=1}^m \pi_i^{x_i} (1 - \pi_i)^{1-x_i}, \quad \boldsymbol{\pi} \in \Theta =]0, 1[^m.$$

The likelihood ratio test is based on the test statistic

$$\Lambda = \frac{\mathbf{L}(\hat{\boldsymbol{\pi}}; \mathbf{X})}{\sup_{\boldsymbol{\pi} \in \Theta} \mathbf{L}(\boldsymbol{\pi}; \mathbf{X})},$$

where $\mathcal{H}_0 : \boldsymbol{\pi} = \hat{\boldsymbol{\pi}}$ is rejected if

$$\Lambda < c$$

for a chosen constant c with $0 < c < 1$. Further,

$$\sup_{\boldsymbol{\pi} \in \Theta} \mathbf{L}(\boldsymbol{\pi}; \mathbf{x}) = \sup_{0 < \pi_1 < 1} \dots \sup_{0 < \pi_m < 1} \prod_{i=1}^m \pi_i^{x_i} (1 - \pi_i)^{1-x_i} = \prod_{i=1}^m 1 = 1, \quad (14)$$

for all $\mathbf{x} \in \{0, 1\}^m$ yields

$$\Lambda = \mathbf{L}(\hat{\boldsymbol{\pi}}; \mathbf{X}).$$

\mathcal{H}_0 is rejected, if the statistic

$$\ln \Lambda = \sum_{i=1}^m \left(X_i \ln \hat{\pi}_i + (1 - X_i) \ln(1 - \hat{\pi}_i) \right) = \sum_{i=1}^m \left(X_i \ln \frac{\hat{\pi}_i}{1 - \hat{\pi}_i} + \ln(1 - \hat{\pi}_i) \right)$$

is less than $\ln(c)$, or equivalently, if the test statistic given in (7) is smaller than a chosen value $k = \ln(c) - \sum_{i=1}^m \ln(1 - \hat{\pi}_i)$.

C Appendix. Proof of Theorem 5.2

The likelihood ratio test statistic is

$$\Lambda = \frac{\sup_{\boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}} \mathbf{L}(\boldsymbol{\pi}; \mathbf{X})}{\sup_{\boldsymbol{\pi} \in \Theta} \mathbf{L}(\boldsymbol{\pi}; \mathbf{X})} = \sup_{\boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}} \mathbf{L}(\boldsymbol{\pi}; \mathbf{X}),$$

²²On the condition that $Y_0 = y_0$.

where the second equality follows from (14). The test statistic T_2 in (9) results from taking the logarithm of

$$\Lambda = \sup_{\boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}} \mathbb{L}(\boldsymbol{\pi}; \mathbf{x}) = \prod_{i=1}^m \sup_{\pi_i \leq \hat{\pi}_i} \pi_i^{x_i} (1 - \pi_i)^{1-x_i} = \prod_{i=1}^m \hat{\pi}_i^{x_i}.$$

Rejecting $\mathcal{H}_0 : \boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}$, if $\Lambda < c$, is equivalent to rejecting \mathcal{H}_0 , if $T_2 < k = \ln c$.

For the proof of the second statement of Theorem 5.2

$$\sup_{\boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}} P_{\boldsymbol{\pi}}(T_2 < k) = P_{\hat{\boldsymbol{\pi}}}(T_2 < k),$$

the test statistic $\Lambda = \prod_{i=1}^m \hat{\pi}_i^{X_i}$ is examined for fixed $\hat{\boldsymbol{\pi}}$ but for alternative distributions of \mathbf{X} with different parameters $\boldsymbol{\pi}$. Let $\Lambda_{\boldsymbol{\pi}} = \prod_{i=1}^m \hat{\pi}_i^{X_i}$ denote the random variable with stochastically independent $X_i \sim \text{Ber}(\pi_i)$ for a parameter vector $\boldsymbol{\pi} = (\pi_1, \dots, \pi_m)$ with $\boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}$. And let $\Lambda_{\hat{\boldsymbol{\pi}}} = \prod_{i=1}^m \hat{\pi}_i^{\hat{X}_i}$ denote the random variable with stochastically independent $\hat{X}_i \sim \text{Ber}(\hat{\pi}_i)$. Then we have $X_i \leq_{st} \hat{X}_i$ for $i = 1, \dots, m$, where \leq_{st} denotes the usual stochastic order for random variables, which is also known as first order stochastic dominance. Since

$$\psi : \mathbb{R}^m \rightarrow \mathbb{R}, \quad \text{with} \quad \psi(a_1, \dots, a_m) = - \prod_{i=1}^m \hat{\pi}_i^{a_i},$$

is a non-decreasing function with respect to all its arguments, Theorem 1.2.4 of Müller & Stoyan (2002, p. 7) can be applied to infer, that $\Lambda_{\hat{\boldsymbol{\pi}}} \leq_{st} \Lambda_{\boldsymbol{\pi}}$, which means

$$P(\Lambda_{\hat{\boldsymbol{\pi}}} \leq t) \geq P(\Lambda_{\boldsymbol{\pi}} \leq t)$$

for all t . Therefore, we have

$$P_{\boldsymbol{\pi}}(\Lambda < c) \leq P_{\hat{\boldsymbol{\pi}}}(\Lambda < c)$$

for all $0 < c < 1$ and for all $\boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}$. Hence,

$$\sup_{\boldsymbol{\pi} \leq \hat{\boldsymbol{\pi}}} P_{\boldsymbol{\pi}}(\Lambda < c) = P_{\hat{\boldsymbol{\pi}}}(\Lambda < c).$$

References

- Basel Committee on Banking Supervision, 2005. Studies on the validation of internal rating systems, Revised Version, May 2005.
- Basel Committee on Banking Supervision, 2006. International convergence of capital measurement and capital standards, A Revised Framework, Comprehensive Version.
- Black, F., Cox, J.C., 1976. Valuing corporate securities: Some effects of bond indenture provisions, *Journal of Finance* 31 (2), 351–367.
- Bluhm, C., Overbeck, L., Wagner, C., 2003. An Introduction to Credit Risk Modeling, Chapman & Hall / CRC Press, Boca Raton.
- Casella, G., Berger, R.L., 2002. Statistical Inference, 2nd edn, Duxbury, Pacific Grove.

- Claessens, S., van Wijnbergen, S., 1993. Secondary market prices and Mexico's Brady deal, *Quarterly Journal of Economics* 108 (4), 965–982.
- Clark, E.A., 1991. *Cross-Border Investment Risk: Applications of Modern Portfolio Theory*, Euromoney Publications, London.
- Crosbie, P., Bohn, J., 2003. *Modeling Default Risk*, Moody's KMV Company.
- Crouhy, M., Galai, D., Mark, R., 2000. A comparative analysis of current credit risk models, *Journal of Banking & Finance* 24 (1-2), 59–117.
- Delianedis, G., Geske, R., 2003. Credit risk and risk neutral default probabilities: Information about rating migrations and defaults, Working Paper.
- Duan, J.-C., 1994. Maximum likelihood estimation using price data of the derivative contract, *Mathematical Finance* 4 (2), 155–167.
- Duan, J.-C., 2000. Correction: Maximum likelihood estimation using price data of the derivative contract, *Mathematical Finance* 10 (4), 461–462.
- Duan, J.-C., Gauthier, G., Simonato, J.-G., 2005. On the equivalence of the KMV and maximum likelihood methods for structural credit risk models, Working Paper, Les Cahiers du GERAD G-2005-06.
- Ericsson, J., Reneby, J., 2005. Estimating structural bond pricing models, *Journal of Business* 78 (2), 707–735.
- Geske, R., 1977. The valuation of corporate liabilities as compound options, *Journal of Financial and Quantitative Analysis* 12 (4), 541–552.
- Geske, R., 1979. The valuation of compound options, *Journal of Financial Economics* 7 (1), 63–81.
- Hull, J.C., 2003. *Options, Futures, and Other Derivatives*, 5th edn, Prentice Hall, Upper Saddle River.
- Jones, E.P., Mason, S.P., Rosenfeld, E., 1984. Contingent claims analysis of corporate capital structures: An empirical investigation, *Journal of Finance* 39 (3), 611–627.
- Keenan, S.C., Carty, L.V., Shtogrin, I., 1998. Historical default rates of corporate bond issuers, 1920 - 1997, Moody's Investors Service, New York.
- Keswani, A., 2000. Estimating a risky term structure of Brady bonds, Lancaster University Management School Working Paper.
- Lehmann, E.L., Romano, J.P., 2005. *Testing Statistical Hypotheses*, 3rd edn, Springer, New York.
- Lehrbass, F.B., 2000. A simple approach to country risk, in: Franke, J., Härdle, W., Stahl, G. (Eds.), *Measuring Risk in Complex Stochastic Systems*, Springer, New York, Berlin, Heidelberg, pp. 33–67.
- Maltritz, D., 2006. Quantifizierung von Souveränrisiken, Dissertation, to appear, Metropolis-Verlag, Marburg.

- Manasse, P., Roubini, N., Schimmelpfennig, A., 2003. Predicting sovereign debt crises, Working Paper.
- Merton, R.C., 1974. On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29 (2), 449–470.
- Müller, A., Stoyan, D., 2002. Comparison Methods for Stochastic Models and Risks, Wiley, Chichester.
- Pesaran, M.H., Timmermann, A., 1992. A simple nonparametric test of predictive performance, *Journal of Business & Economic Statistics* 10 (4), 461–465.
- Petrov, V.V., 1995. Limit Theorems of Probability Theory: Sequences of Independent Random Variables, Clarendon Press, Oxford.
- Ronn, I.E., Verma, A.K., 1986. Pricing risk-adjusted deposit insurance: An option-based model, *Journal of Finance* 41 (4), 871–895.

Dresden Discussion Paper Series in Economics

- 12/04 **Broll, Udo / Hansen, Sabine / Marjit, Sugata:** Domestic labor, foreign capital and national welfare
- 13/04 **Nyamtseren, Lhamsuren:** Challenges and Opportunities of Small Countries for Integration into the Global Economy, as a Case of Mongolia
- 01/05 **Schubert, Stefan / Broll, Udo:** Dynamic Hedging of Real Wealth Risk
- 02/05 **Günther, Edeltraud / Lehmann-Waffenschmidt, Marco:** Deceleration - Revealed Preference in Society and Win-Win-Strategy for Sustainable Management. Concept and Experimental Evidence
- 03/05 **Sennewald, Ken:** Controlled Stochastic Differential Equations under Poisson Uncertainty and with Unbounded Utility
- 04/05 **Sennewald, Ken / Wälde, Klaus:** "Itô's Lemma" and the Bellman equation: An applied view
- 05/05 **Neumann, Anne / Siliverstovs, Boriss:** Convergence of European Spot Market Prices for Natural Gas?
- 06/05 **Hirschhausen, Christian von / Cullmann, Astrid:** Efficiency Analysis of German Electricity Distribution Utilities
- 07/05 **Seitz, Helmut / Freigang, Dirk / Kempkes, Gerhard:** Demographic Change and Federal Systems: Some Preliminary Results for Germany
- 08/05 **Bemmann, Martin:** Verbesserung der Vergleichbarkeit von Schätzgüteregebnissen von Insolvenzprognosestudien
- 09/05 **Thum, Marcel:** Korruption und Schattenwirtschaft
- 10/05 **Seitz, Helmut / Kempkes, Gerhard:** Fiscal Federalism and Demography
- 01/06 **Bieta, Volker / Broll, Udo / Milde, Hellmuth / Siebe, Wilfried:** A Strategic Approach to Financial Options
- 02/06 **Battermann, Harald L. / Broll, Udo:** Utility Functions of Equivalent Form and the Effect of Parameter Changes on Optimum Decision Making
- 03/06 **Broll, Udo / Wahl, Jack E.:** Value at risk, Equity and Diversification
- 04/06 **Broll, Udo / Fuchs, Frank:** Optimale Fakturierung im Außenhandel
- 05/06 **Pontes, Jose Pedro:** A Non-monotonic Relationship between FDI and Trade
- 06/06 **Lehmann-Waffenschmidt, Marco:** Self-Referential Optimal Advising When Reactions are Delayed
- 07/06 **Dittrich, Marcus / Schirwitz, Beate:** A Dynamic Model of Union Behaviour. The Role of an Endogenous Outside Option and Bargaining Centralisation
- 08/06 **Kempkes, Gerhard / Pohl, Carsten:** The Efficiency of German Universities – Some Evidence from Non-Parametric and Parametric Methods
- 09/06 **Heinzel, Christoph / Winkler, Ralph:** Gradual versus structural technological change in the transition to a low-emission energy industry - How time-to-build and differing social and individual discount rates influence environmental and technology policies
- 10/06 **Heinzel, Christoph:** Schumpeter and Georgescu-Roegen on the foundations of an evolutionary analysis – The problem of qualitative change, its methodical implications and analytical treatment
- 11/06 **Broll, Udo / Wahl, Jack E.:** Bankmanagement mit Value at Risk
- 12/06 **Karmann, Alexander / Huschens, Stefan / Maltritz, Dominik / Vogl, Konstantin:** Country Default Probabilities: Assessing and Backtesting

