Default probability estimation in small samples - with an application to sovereign bonds

Orth, Walter

University of Cologne

28 September 2011

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 5/11

Default Probability Estimation in Small Samples – With an Application to Sovereign Bonds

by

Walter Orth

First draft: September 28, 2011
This draft: February 9, 2012

DISKUSSIONSBEITRÄGE ZUR STATISTIK UND ÖKONOMETRIE

SEMINAR FÜR WIRTSCHAFTS- UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN

Albertus-Magnus-Platz, D-50923 Köln, Deutschland

Default Probability Estimation in Small Samples – With an Application to Sovereign Bonds

Walter Orth*

Abstract

In small samples and especially in the case of small true default probabilities, standard approaches to credit default probability estimation have certain drawbacks. Most importantly, standard estimators tend to underestimate the true default probability which is clearly an undesirable property from the perspective of prudent risk management. As an alternative, we present an empirical Bayes approach to default probability estimation and apply the estimator - which is capable of multi-period predictions - to a comprehensive sample of Standard & Poor's rated sovereign bonds. By means of a simulation study, we then show that the empirical Bayes estimator is more conservative *and* more precise under realistic data generating processes.

JEL classification: C11, C41, G15, G28

Keywords: Low-default portfolios, empirical Bayes, sovereign default risk, simulation study

*Department of Statistics and Econometrics, University of Cologne; walter.orth@uni-koeln.de

1. Introduction

We consider the estimation of credit default probabilities for obligors with given rating grades. Assigning default probabilities (PDs) to rating grades is clearly of high relevance for the purpose of financial risk management and the Basel II Accord has considerably increased the attention for this task. In this article, our focus is on problems that arise when sample sizes are rather small. Especially for higher grades, typical samples often contain only a few, if any, default events. Such portfolios are often referred to as low-default portfolios in the related literature. Important examples for such sparse datasets are samples of sovereigns or financial institutions. In a survey of UK banks, "insufficient default data" for PD estimation are reported by 90% of the banks with respect to sovereign obligors while the same figure is 62% for obligors being banks as well (BBA, 2004).

Standard approaches applied to small samples and especially to samples with few defaults have serious drawbacks. Besides the obvious effect that estimation uncertainty is high the skewness of the sampling distribution leads to a high probability of underestimating the true PD. Intuitively, given small true PDs and small sample sizes, it is quite likely not to observe any default event in a particular sample leading to a PD estimate of zero under standard approaches. The likelihood of underestimating the true default probability rises as i) the true PD decreases, ii) the sample size decreases and iii) the correlation of default events increases (Kurbat & Korablev, 2002).[1] Given these properties, it would be desirable to improve upon standard estimators by applying a more efficient and more conservative estimator in small samples. In fact, in the context of Basel II, a conservative approach is demanded from the regulatory side. In § 416 of Basel II (BCBS, 2006a), it is stated that "...where limited data are available, a bank must adopt a conservative bias to its analysis." Similarly, from § 451: "In order to avoid over-optimism, a bank must add to its estimates a margin of conservatism that is related to the likely range of errors." A certain degree of conservatism may be seen as a general principle of prudent risk management but we will further show in section 5 that, given estimation uncertainty, an upward bias of the PD estimator is necessary to estimate Basel II capital requirements unbiasedly.

PD estimates for a one-year prediction horizon are particularly relevant as inputs for the calculation of Basel II capital requirements. However, the scope of this article is somewhat broader in the sense that we will consider a flexible multi-period prediction horizon. Given the multi-period risk usually faced by lenders, a multi-period time horizon is naturally of interest and its importance may further rise due to plans of the International Accounting Standards Board and the Basel Committee on Banking Supervision to base loss provisions upon the expected loss over the whole life of the credit portfolio (BCBS, 2009). But even for a one-year prediction horizon a multi-period view is useful since it allows using, say, quarterly or monthly data and thus the information of within-year rating changes and

[1] We will provide numerical evidence on the probability to underestimate the true PD in our simulation study in section 5.

censoring events.[2] This should be of particular importance in the case of small samples. Nevertheless, in this paper we present results that may also be interesting for the reader who is primarily interested in a single-period view since this is just a special case of our more general approach.

The problem of low-default portfolios has already received some attention in the literature. One approach is to employ the idea of confidence intervals and to use an appropriate upper confidence bound as a conservative PD estimator (Pluto & Tasche, 2006; Benjamin et al., 2006). However, these studies consider a fixed one-year prediction horizon and do not use potentially available within-year information. As will be discussed in the upcoming section, the confidence bound approach is not easily generalized to our more flexible multi-period setup. Further, confidence intervals for realistic sample sizes are often wide (Lawrenz, 2008) and would therefore often result in a sharp increase of PD estimates. It is then doubtful if a bank would be willing to use such an estimate for internal uses such as pricing, something that is explicitly intended within the Basel II framework and also monitored by the so-called IRB Use Test (BCBS, 2006b).

Alternatively, Kiefer (2009) suggests a Bayesian approach to PD estimation. In the Bayesian framework, the standard estimator is shrunk towards some prior guess which potentially leads to more conservative and more accurate estimates. A crucial part of a Bayesian analysis is the choice of the prior distribution. Kiefer (2009) proposes to specify the prior by expert elicitation while in a recent study Tasche (2011) proposes to use non-informative priors. In this paper, we present an empirical Bayes approach where the prior distribution is estimated from data. Our estimator has a closed-form representation. We apply our estimator to a comprehensive dataset of Standard & Poor's rated sovereigns and study its properties in detail by means of a simulation study.

We do not consider models for rating migrations although these could also be used for PD estimation. For instance, the use of Markovian rating migration models is quite standard and has the benefit that it can remove PD estimates of zero if the time intervals are chosen small enough (Lando & Skodeberg, 2002). However, there is strong evidence against the Markovian assumption as migration probabilities have been found to depend on the direction of the prior rating action for corporates (Lando & Skodeberg, 2002) and for sovereigns as well (Fuertes & Kalotychou, 2007). In particular, downgrades are more often followed by subsequent downgrades than implied by a Markovian model so that PDs derived from Markovian migration models tend to be downward biased (Hanson & Schuermann, 2006).[3] More sophisticated models like Hidden Markov Models (Christensen et al., 2004) have been proposed in the literature but come at the cost of considerably

[2] For instance, when a firm was downgraded from, say, rating A to rating B within a year and subsequently defaulted in this year, an approach that uses only year-end data misses the information that a B rated firm has defaulted.

[3] Since the cited empirical evidence concerns agency ratings it is of course possible that for any internal rating system a Markovian approach is appropriate. Still, one has to be aware that any violation of the Markovian assumption can lead to biased estimates.

more complexity not least in terms of a higher number of parameters. In small samples, such an augmented parameterization is likely to cause instability, i.e. high variance of the parameter estimates and is thus rather not suitable in our context.

The rest of the paper is organized as follows. In the next section, we will introduce our notation and present what we will call the standard estimator. In section 3, we will present the empirical Bayes approach, which will then be applied - together with the standard estimator - to our sovereign data in section 4. Section 5 contains the simulation study which evaluates and compares both the standard estimator and the empirical Bayes estimator while section 6 concludes.

2. Standard default probability estimation

The standard estimator that we will present in this section is the approach used by the major rating agencies in their calculation of cumulative default rates (Hamilton & Cantor, 2006) which are the counterparts to PDs.[4] Let us first introduce the notation. All the obligors that have the rating r at time t, $t = 1, \ldots, T$, form a so-called cohort. We denote by $N^r_{t,1}$ the number of obligors that comprise the cohort at its beginning (t) and we denote by $N^r_{t,s}$ those members of the cohort that are still at risk before period $t+s$. Being at risk means that an obligor has not defaulted or is not censored in the first $s-1$ periods after the cohort building date t.[5] Out of the $N^r_{t,s}$ obligors entering period $t+s$, let $D^r_{t,s}$ be the number of those that default in period $t+s$ and let $L^r_{t,s}$ be the number of those which are lost, i.e. which are censored, in period $t+s$. Further, let λ^r_s be the marginal default rate or discrete-time hazard rate which is the probability that an obligor rated r at a certain point in time will default s periods later conditional on surviving the first $s-1$ periods. If we define Y_{it} to be the discretely measured lifetime (the time until default) of obligor i that starts in period t and define R_{it} to be the corresponding rating, we can write this probability formally as

$$\lambda^r_s = P(Y_{it} = s | Y_{it} > s-1, R_{it} = r)\,. \tag{1}$$

Under our notation, the standard estimator of the marginal default rate is

$$\widehat{\lambda}^r_s = \frac{\sum_{t=1}^{T} D^r_{t,s}}{\sum_{t=1}^{T} N^r_{t,s} - L^r_{t,s}/2}\,. \tag{2}$$

Usually the main interest is on the corresponding cumulative default rates or default probabilities which we will denote by PD^r_s and define as

$$PD^r_s = P(Y_{it} \leq s | R_{it} = r)\,. \tag{3}$$

[4] It is also largely equal to the approach of Altman (1989). However, in that study the analysis is restricted to the cumulative default rates of newly issued bonds.

[5] Censoring means that the obligor is no longer observed - without defaulting - after some of the $s-1$ periods. Censored data are quite common and occur, for instance, due to withdrawals of ratings, the end of the sample period or gaps in the dataset.

The estimator for the cumulative default rates is easily constructed from the marginal default rate estimators:

$$\widehat{PD}_s^r = 1 - \prod_{j=1}^{s}(1 - \widehat{\lambda}_j^r) \tag{4}$$

Let us briefly interpret what the estimator actually does. The estimator starts with the estimation of marginal default rates by taking a weighted average of the marginal default rates of individual cohorts, $\frac{D_{t,s}^r}{N_{t,s}^r - L_{t,s}^r/2}$. The weights are here the adjusted number of obligors at risk, $N_{t,s}^r - L_{t,s}^r/2$. The adjustment involves the subtraction of half of the censored observations being based on the assumption that censored obligors have still survived on average half of the corresponding period. Then, the estimator constructs estimators for the cumulative default rate from the marginal default rate estimates. While it would be also possible to take a weighted average of the cumulative default rates of individual cohorts, averaging marginal default rates results in more efficient estimation as was already shown by Cutler & Ederer (1958). The presented estimator is also known as the life-table or actuarial estimator and is approximately equal to the widely-used Kaplan-Meier or Product-Limit estimator as the period length becomes small. The Product-Limit estimator has been shown to be a Nonparametric Maximum Likelihood estimator (Johansen, 1978) but note that in our setting we rather have a pseudo ML estimator since our observations are not independent. To see this, notice that the same obligor enters a new cohort every period and that default events are used several times in the estimation process. For instance, consider an obligor which is rated A in period t, stays rated A in period $t+1$ and subsequently defaults in period $t+2$. The same default event enters the calculation of $\widehat{\lambda}_1^A$ and $\widehat{\lambda}_2^A$. Put another way, the analysis is based on a sequence of overlapping lifetimes, $Y_{it}, Y_{i,t+1}, \ldots$, which clearly leads to dependencies. As the simulation study in section 5 will confirm, these dependencies (and additional dependencies through common shocks) do not introduce any relevant bias to the estimator. However, conventional formulas for confidence intervals will be misleading with standard errors ignoring the dependence usually being downward biased. This is why it is not easily possible to generalize the confidence bound approach mentioned in the introduction to a setup with a multi-period prediction horizon. A valid confidence bound approach for multi-period PDs would have to take into account dependence through overlapping lifetimes *and* through common shocks which is likely to be very challenging. Note that if a one-period view is employed, i.e. only $\lambda_1^r = PD_1^r$ is estimated, dependencies through overlapping lifetimes vanish and the problem gets more tractable.[6]

What remains to be specified for an empirical analysis is the period length. Rating agencies differ in this respect. While Moody's has switched to building cohorts on a monthly basis since 2005 (Hamilton & Cantor, 2006), Standard & Poor's uses cohorts of obligors built at the end of every calendar year (Standard & Poor's, 2011a). In our

[6]Pluto & Tasche (2006) show how to calculate confidence bounds for one-period PDs in the presence of common shocks. With T sample periods, this involves T-dimensional integration.

application, we will use monthly cohorts and accordingly construct our PD estimates using monthly hazard rates in order to use as much sample information as possible. Since more than one rating change per month for the same obligor occurs only very rarely, almost no information is lost under a monthly periodicity.

3. Empirical Bayes estimation

Bayesian parameter estimation is a potentially very useful approach especially in the case of small samples. The reason for this is that in small samples the data provide only little information about the parameter of interest so that the incorporation of prior information can be particularly helpful. To use such prior information, a Bayesian data analyst specifies a prior distribution of the parameter of interest. This can be done by different means. Possibilities that have been proposed in the context of PD estimation include the specification of the prior by means of expert elicitation (Kiefer, 2009) or the use of uninformed priors (Tasche, 2011). In this paper, we will alternatively propose a data-driven way to specify the prior distribution by using an empirical Bayes (EB) approach.[7] To enable the estimation of the prior distribution, further datasets besides the original sample are needed. For instance, in our empirical analysis regarding sovereigns, we will further use data on firms to estimate the prior distribution. In many practical situations, such auxiliary datasets will be available. For example, a bank will typically have a variety of different portfolios and the information from all these portfolios can be used within the EB approach to estimate the PDs for each particular portfolio. Similarly, default histories referring to external ratings can be used as a prior for PD estimation based on internal data. The combination of different datasets is - without any explicit proposal - also mentioned from the regulatory side as one tool for PD estimation in the case of low-default portfolios (BCBS, 2005).

We will now formally introduce our EB estimator and subsequently give some further discussion. Like in the case of the standard estimator, we will start with the estimation of marginal default rates which are then used to construct cumulative default rates. Suppose that we have G different groups or portfolios where $G \geq 2$. We make the parametric assumption that for each group g, $g = 1, \ldots, G$, the marginal default rates are a priori beta distributed,

$$\lambda_s^{r,g} \sim beta(\alpha_s^r, \beta_s^r)\,. \tag{5}$$

Note that each group has the same prior parameters. Further, we assume that the conditional distribution of the number of defaults in period s is binomial,

$$D_s^{r,g} | \lambda_s^{r,g} \sim Bin(\widetilde{N}_s^{r,g}, \lambda_s^{r,g})\,, \tag{6}$$

where we have now, to simplify notation, skipped the index t to indicate aggregation over the cohort building dates, i.e. $D_s^{r,g} = \sum_{t=1}^{T} D_{t,s}^{r,g}$ and $\widetilde{N}_s^{r,g} = \sum_{t=1}^{T}(N_{t,s}^{r,g} - L_{t,s}^{r,g}/2)$. The

[7] For an introduction to EB methods we recommend Casella (1985) and Carlin & Louis (2008).

presented framework is known as the beta-binomial model and is quite common for the Bayesian analysis of proportions. The beta distribution is a pretty flexible distribution for parameters bounded in the interval [0,1] and has also been suggested by Kiefer (2009).[8] The crucial part of the binomial assumption is the conditional independence of default events. Note that although we aggregate over different cohort building dates we do not use the same default event more than once for fixed s. However, we disregard the dependence of default events induced by common shocks to keep the analysis as simple as sensibly possible. In our simulation study of section 5, we will show that the estimator works well even for data generating processes that involve dependencies through common shocks.[9]

The next step is now to estimate the prior parameters. We do so by using the Method of Moments hereby essentially following the analysis of Kleinman (1973). For convenience, we reparameterize the beta distribution setting $\mu_s^r = \alpha_s^r/(\alpha_s^r + \beta_s^r)$ to be the prior mean of $\lambda_s^{r,g}$ and $\tau_s^r = 1/(1 + \alpha_s^r + \beta_s^r)$ to be a measure of prior precision. We estimate μ_s^r as a weighted average of the group-specific standard marginal default rate estimates:

$$\widehat{\mu}_s^r = \sum_{g=1}^{G} w_s^{r,g} \frac{D_s^{r,g}}{\widetilde{N}_s^{r,g}} = \sum_{g=1}^{G} w_s^{r,g} \widehat{\lambda}_s^{r,g} \tag{7}$$

The formula we use to estimate τ_s^r is

$$\widehat{\tau}_s^r = \frac{\frac{G-1}{G} \sum_{g=1}^{G} w_s^{r,g} (\widehat{\lambda}_s^{r,g} - \widehat{\mu}_s^r)^2 - \widehat{\mu}_s^r (1 - \widehat{\mu}_s^r) \left(\sum_{g=1}^{G} w_s^{r,g} (1 - w_s^{r,g}) / \widetilde{N}_s^{r,g} \right)}{\widehat{\mu}_s^r (1 - \widehat{\mu}_s^r) \left(\sum_{g=1}^{G} (1 - 1/\widetilde{N}_s^{r,g}) w_s^{r,g} (1 - w_s^{r,g}) \right)}. \tag{8}$$

See Kleinman (1973) for a detailed derivation of these formulas. Natural choices for the weights are equal weights, i.e. $w_s^{r,g} = 1/G$, or the number of observations for each group so that $w_s^{r,g} = \widetilde{N}_s^{r,g} / \sum_{g=1}^{G} \widetilde{N}_s^{r,g}$. Kleinman (1973) shows that the optimal weights depend on the true parameters and thus proposes to use one iteration to refine the estimates, namely to set

$$w_s^{r,g} = \frac{\widetilde{N}_s^{r,g}}{1 + \widehat{\tau}_s^r (\widetilde{N}_s^{r,g} - 1)} \Bigg/ \sum_{j=1}^{G} \frac{\widetilde{N}_s^{r,j}}{1 + \widehat{\tau}_s^r (\widetilde{N}_s^{r,j} - 1)}, \tag{9}$$

using a preliminary estimate of τ_s^r to get improved weights which are subsequently employed to re-estimate the prior parameters. In our implementation, we used this one-time iteration step with starting weights $w_s^{r,g} = 1/G$. However, we also experimented with an omission of the iteration step and found no large sensitivity of the results in this respect. Since there is no guarantee that $\widehat{\tau}_s^r$ will be in the interval $(0, 1)$ which is necessary for a proper prior the estimates of $\widehat{\tau}_s^r$ should be truncated at zero and one, respectively.

[8] Kiefer (2010) shows how the beta distribution can be generalized for even more flexibility. Note that prior distributions with more parameters will increase the minimum number of groups.

[9] The incorporation of dependence as an extension of our estimator seems to be interesting and could be done along the lines of Kiefer (2010) and Tasche (2011). However, especially with multiple periods of data, the computational burden will be considerable making a simulation study as the one in section 5 hardly affordable.

With the estimated prior parameters at hand, we can apply the Bayesian theorem to arrive at the posterior distribution of our parameters. Since the beta distribution is the conjugate prior for the binomial distribution, the posterior distribution of $\lambda_s^{r,g}$ is beta as well. The mean of the posterior distribution minimizes the Bayes risk under quadratic loss functions and is the standard choice for a Bayesian point estimator. In our case, the posterior mean, i.e. our EB estimator for $\lambda_s^{r,g}$, can be written as

$$\widehat{\lambda}_{s,EB}^{r,g} = \frac{1-\widehat{\tau}_s^r}{1+\widehat{\tau}_s^r(\widetilde{N}_s^{r,g}-1)}\widehat{\mu}_s^r + \frac{\widehat{\tau}_s^r\widetilde{N}_s^{r,g}}{1+\widehat{\tau}_s^r(\widetilde{N}_s^{r,g}-1)}\widehat{\lambda}_s^{r,g}\,. \tag{10}$$

The EB estimator is obviously a weighted average of the prior mean (which itself is a weighted average of the group-specific standard estimates) and the standard marginal default rate estimate for group g. The estimator can be interpreted as a shrinkage estimator since it shrinks the standard estimates, $\widehat{\lambda}_s^{r,g}$, towards the prior means which are equal for all groups. Note also that the weighting scheme is such that there is a smooth transition to the standard estimator if $\widetilde{N}_s^{r,g}$ grows. Thus, the amount of shrinkage will, in line with intuition, decline as the sample size of a specific group increases. We provide R code for Formulas (7)-(10) in appendix A.

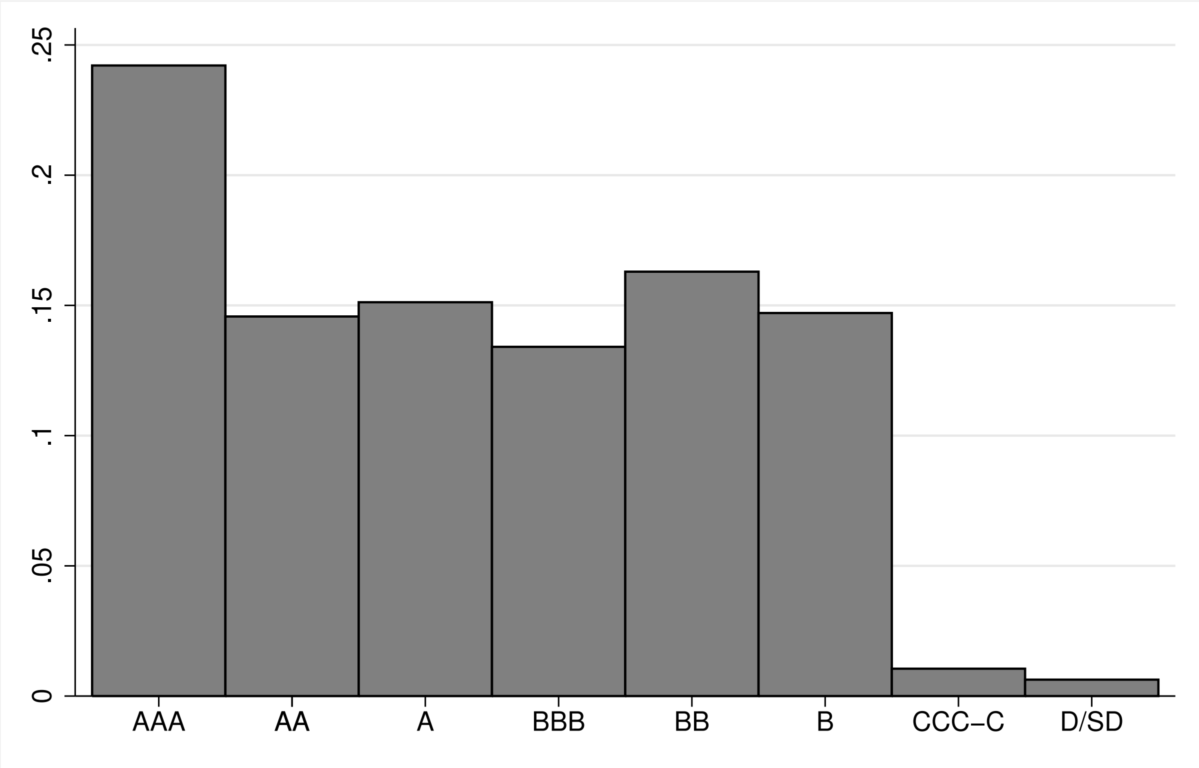
Like in the preceding section, our estimate for the PD is constructed from the estimated marginal default rates,

$$\widehat{PD}_{s,EB}^{r,g} = 1 - \prod_{j=1}^{s}(1-\widehat{\lambda}_{j,EB}^{r,g}) \tag{11}$$

It is worth mentioning that our estimator minimizes the Bayes risk with respect to the marginal default rates instead of the cumulative default rates. Doing the latter would actually be preferable but would considerably increase the complexity of the problem in our setting since we would have to deal with the dependencies due to overlapping lifetimes. Note, however, that our estimator can be interpreted as minimizing the Bayes risk for the PD under a working independence assumption. In this case, our estimator is equal to the one derived by Hjort (1990) for the discrete-time case.

Although derived from Bayesian theory, EB methods have been shown to be an improvement over standard Maximum Likelihood methods in many applications even by frequentist measures such as Mean Squared Error (Casella, 1985). Usually EB methods reduce the variance of an estimator at the cost of introducing some bias. Consider for instance our application where we will combine sovereign and corporate datasets. The more the true PDs for both groups are apart the larger will be the bias introduced by the EB approach. However, if the differences are rather small the effect of variance reduction will prevail and lead to smaller Mean Squared Errors. Especially in small samples, where the variance of the standard estimator is high, the potential gains from variance reduction can be substantial. In our simulation study in section 5, we will illustrate this bias-variance trade-off under realistic scenarios. Moreover, if conservativeness is by itself desirable, as it is stated at least by regulators with respect to PD estimation, a moderate upward bias induced by EB methods may even be seen as a benefit rather than a weakness.

Figure 1: Rating distribution of S&P rated sovereign bonds, 1975-2011



4. Application to sovereign bonds

Sovereign bonds provide possibly the most important application for our methods since they are among the most important asset classes and sovereign defaults are rare events. In this study, we use the complete rating and default histories of sovereigns with public ratings from Standard & Poor's in the period from January 1975 until April 2011.[10] The data are from Standard & Poor's (2011b) and consist of 130 sovereigns observed over a total of 23,014 country-months. The dataset may thus appear not that small but the sample size per rating class is of course considerably lower and, importantly, we observe only 15 default events. More precisely, these default events are foreign-currency selective defaults. Accordingly, we will use foreign-currency issuer credit ratings in our analysis since these have longer rating histories and are probably in most cases more relevant to investors than local-currency ratings. Figure 1 shows the rating distribution in our sample. Apart from the data concerning sovereigns, we further utilize S&P rating and default histories of North American public firms from Compustat covering the period from January 1981 until April 2011. The corresponding ratings used here are S&P's long term issuer credit ratings and are on the same scale and have the same definition as their sovereign counterparts. The latter can be seen as a justification of our Bayesian assumption that there is no difference between sovereigns and firms *a priori*. This second

[10]Standard & Poor's also rates certain sovereigns on a confidential basis. See Standard & Poor's (2011a).

Table 1: Cumulative default rate estimates (1-10 years)

Panel A: Empirical Bayes estimator for sovereigns

	1	2	3	4	5	6	7	8	9	10
AAA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AA	0.00	0.01	0.05	0.08	0.08	0.09	0.12	0.13	0.17	0.22
A	0.06	0.16	0.33	0.50	0.65	0.79	0.92	1.08	1.26	1.42
BBB	0.20	0.57	1.07	1.65	2.24	2.81	3.40	3.95	4.60	5.34
BB	0.76	2.42	4.24	5.70	7.56	9.40	11.27	13.04	14.74	15.95
B	3.88	7.64	11.43	15.80	19.58	23.07	26.01	28.66	31.04	33.21
CCC-C	24.38	33.66	39.87	43.89	48.02	52.72	66.77	82.27	82.46	82.88

Panel B: Standard estimator for sovereigns

	1	2	3	4	5	6	7	8	9	10
AAA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BBB	0.00	0.66	1.60	2.62	3.71	4.85	5.27	5.27	5.27	5.27
BB	0.56	1.88	3.06	3.66	5.10	6.75	8.74	10.29	11.23	11.43
B	2.60	5.16	7.59	11.01	13.12	15.48	18.30	22.17	25.44	29.02
CCC-C	32.27	44.50	51.56	55.76	63.45	72.19	83.85	91.92	91.92	91.92

Panel C: Standard estimator for corporates

	1	2	3	4	5	6	7	8	9	10
AAA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AA	0.00	0.01	0.05	0.08	0.09	0.10	0.13	0.14	0.18	0.24
A	0.06	0.16	0.33	0.51	0.67	0.81	0.95	1.11	1.29	1.45
BBB	0.20	0.57	1.06	1.63	2.20	2.76	3.36	3.92	4.59	5.36
BB	0.77	2.45	4.29	6.15	8.02	9.86	11.72	13.50	15.24	16.89
B	4.47	9.96	14.91	19.21	22.98	26.38	29.22	31.70	33.96	35.97
CCC-C	24.20	33.31	39.49	43.52	47.48	50.60	52.89	53.58	54.07	55.17

dataset is large (5355 firms, 563,809 firm-months, 755 defaults) and will be used for the EB approach as explained in the previous section.

Table 1 shows cumulative default rate estimates for sovereigns using the EB approach in Panel A and under the standard estimator in Panel B. Further, Panel C shows the standard estimator applied to our corporate data.[11] To start with Panel B, we see that we get PD estimates of zero throughout all time horizons for the three highest grades and under a one-year horizon even for BBB rated sovereigns. This is clearly an unpleasant feature since such estimates are anticonservative and also not in line with market perceptions given that credit default swaps are traded even for highly rated sovereigns. In contrast, the EB estimator manages to remove most of the zeroes. Due to the relatively small

[11]The EB estimator for corporates is, as expected from theory, very close to the standard estimator and thus not reported here.

size of the sovereign sample the EB estimator is dominated by the standard estimator for corporates as can be seen by comparing Panel A and Panel C. However, this closeness is varying. For instance, in the case of sovereigns rated B, where we have relatively much information in the sense that we have some defaults and not too few sovereigns rated B, we observe that the sovereign estimates are less close to the corporate estimates as they are for other grades. Overall, we see that the sovereign PD estimates are more conservative under the EB approach while the increase in conservativeness seems to be at a reasonable degree.

We now go on to analyze the economic impact of our different estimators. The two applications we have chosen are the estimation of expected returns and the calculation of economic capital. With respect to the former, we consider sovereign bond investments with a maturity of up to 10 years. For these bonds, we consider a simple hold-to-maturity strategy and calculate expected returns by replacing contractual cash flows with their expected values and computing the corresponding yield-to-maturity. Besides a term structure of PDs, this requires an assumption for the recovery rate, i.e. the proportion of the face value of the bond that is recovered if default occurs. For our calculations, we assume a recovery rate of 0.55 which is the middle of the interval [0.5, 0.6] reported in Standard & Poor's (2011c) as the estimated historical average sovereign recovery rate. Our choice for the recovery rate also coincides with the loss given default (= 1− recovery rate) assumption of 0.45 which is prescribed in the foundation Internal Ratings Based (IRB) approach of Basel II (BCBS, 2006a, §287). For the results shown in Table 2, we have selected one USD denominated bond for each rating grade with the exception of the CCC-C grades since no sovereign had such a rating on May 2, 2011 (the hypothetical bond purchase date). By comparing the maximum return, i.e. the return that an investor will receive if no default occurs, with the expected returns we can see which part of the maximum return an investor can expect to lose on average by taking the risk of the corresponding bond investment. Under each estimator, the results are in line with the basic risk-return paradigm in the sense that expected returns monotonically rise as ratings worsen. Further, with the exception of the BBB class, the reward for risk is estimated to be lower under the EB approach which, of course, directly follows from its relative conservatism.

Our second application is to use our one-year PD estimates as inputs to the Basel II capital formula (BCBS, 2006a, §272).[12] Besides the PD, the formula requires as an input an estimate of the loss given default which we again set to 0.45. The correlations which are also part of the formula are defined by the regulators as a function of the PD. Further, under the advanced IRB approach there are potential maturity adjustments if the effective maturity differs from 2.5 years. To facilitate comparisons, we assume the standard maturity of 2.5 years. Table 3 shows the results for the capital requirements. The

[12] Under the new regulatory initiative called Basel III the capital formula is not intended to be changed in its structure. However, the capital requirements are likely to be scaled up. See BCBS (2010) for details.

Table 2: Expected returns for selected USD denominated sovereign bonds

Country	Rating	Maturity	Max. return	Expected returns: Standard estimator	Expected returns: Empirical Bayes
USA	AAA	2/2021	3.363	3.363	3.363
Qatar	AA	1/2020	4.584	4.584	4.576
Poland	A	7/2019	4.750	4.750	4.687
Lithuania	BBB	3/2021	5.580	5.312	5.335
Egypt	BB	4/2020	6.557	5.962	5.766
Argentina	B	6/2017	8.783	7.332	6.982

Sovereign bond data are from Boerse Frankfurt. Expected returns are calculated under the assumption of a bond purchase on May 2, 2011, and a hold-to-maturity strategy.

Table 3: Basel II capital requirements

	ST	SE	EB	EB*	% of MP
AAA	0.00	0.00	0.00	0.00	46.6
AA	0.00	0.00	0.00	0.77	35.7
A	1.60	0.00	1.69	1.69	8.4
BBB	4.00	0.00	3.48	3.48	5.8
BB	8.00	5.88	6.66	6.66	2.8
B	8.00	9.87	11.06	11.06	0.7
CCC-C	12.00	19.85	19.67	19.67	0.0
MP	0.65	0.23	0.61	0.88	

Numbers in percentage points. ST: Basel II Standardised Approach; SE: Standard estimator; EB: Empirical Bayes estimator; EB*: Empirical Bayes estimator with one-year AA PD calculated by scaling down the associated three-year PD. MP refers to an approximate market portfolio (see the text for details).

first column contains the capital requirements under the standardised approach of Basel II where banks do not estimate PDs themselves and use instead fixed ratings-based risk weightings (BCBS, 2006a, §53). The corresponding capital requirements are intended to be conservative in order to give banks an incentive to intensify their own risk analysis and to move eventually to the IRB approach. Comparing the first with the second column, we see that indeed under an IRB approach that uses our standard estimator the capital requirements are considerably lower than under the standardised approach. However, the capital ratios for this case seem to be very liberal since no capital at all is needed for BBB or better rated sovereigns including, for instance, Portugal and Ireland at the end of our sample period. In contrast, the EB estimates are more conservative thereby not implying unrealistically high levels of capital as is seen by their closeness to the standardised approach. In the column of Table 3 which is labelled by EB*, we have kept the original

EB estimates with the exception of the AA category. To get a non-zero estimate for this class, we used a proposal from BCBS (2005) to scale down multi-year PDs in the case of sparse data. Specifically, we used the non-zero AA three-year PD estimated under the EB approach to calculate the one-year PD under the assumption of constant marginal default rates, i.e. $\widehat{PD}_1^{AA,*} = 1-(1-\widehat{PD}_3^{AA})^{1/3} = 0.015\%$. At first sight, the new AA capital ratio of 0.77% seems to be negligibly small. However, matters change if we analyze the impact on the capital requirements of a bank which holds an approximate market portfolio. The composition of the market portfolio is given in the last column of Table 3 and is calculated from data of the Bank for International Settlements on the total amounts of outstanding debt in government securities.[13] Holding the market portfolio, a bank would increase its capital requirement to 0.88% from 0.61% under the unadjusted EB approach and from 0.23% under the standard estimator. These differences are of course substantial and show the high sensitivity of capital ratios to PD estimates. Moreover, it is obvious that the standard PD estimator is far from being conservative in this respect.

5. Simulation study

While we have seen that EB estimators have nice theoretical properties and give reasonable results in our empirical application it is clearly of interest to study the performance of the EB estimator and the standard estimator in more detail. Out-of-sample tests are no appropriate option in our case since our small sample size would not allow us to draw meaningful conclusions. Instead, we will evaluate our estimators by means of a simulation study. The specification of our data generating process is as follows. With respect to the sample size, we stick to the data used in the previous section. More precisely, we drop all observations from our datasets with the exception of the observations where a firm or sovereign first entered the dataset. For instance, the United States enter our dataset in January 1975 with a AAA rating which remains constant until the end of our sample. For our simulation, we keep only the rating in January 1975 whereas the subsequent ratings are now filled up in the simulation process.[14] We choose a Markovian rating migration model (again on a monthly basis) to simulate rating transitions. While we have argued that the Markovian model has serious drawbacks for PD estimation it should nevertheless be suitable for our simulations since none of our estimators relies on the Markovian assumption. We tested several structures for the data-generating migration

[13] The data are available under http://bis.org/statistics/secstats.htm. We aggregated the outstanding amounts of international and domestic debt securities per government (Tables 12D and 16A) as of December 2010. The corresponding S&P ratings for the same date were used to compute aggregated amounts of debt per rating class. Note that no government was rated CCC or lower at this point in time so that the CCC-C category estimates do not influence the market portfolio calculations.

[14] In our corporate sample, we have firms that are not observed for some periods and then return at a later point in time. To account for these censoring events, we keep the first rating of these firms after their return for our simulations and treat them as if they were new firms.

matrix and found the following one to lead to a realistic migration behavior as well as reasonable levels of pseudo-true PDs:[15]

	AAA	AA	A	BBB	BB	B	$CCC-C$	D/SD
AAA	$1-\frac{7}{4}m$	m	$m/2$	$m/4$	0	0	0	0
AA	m	$1-\frac{11}{4}m$	m	$m/2$	$m/4$	0	0	0
A	$m/2$	m	$1-\frac{13}{4}m$	m	$m/2$	$m/4$	0	0
BBB	$m/4$	$m/2$	m	$1-\frac{7}{2}m$	m	$m/2$	$m/4$	0
BB	0	$m/4$	$m/2$	m	$1-\frac{7}{2}m$	m	$m/2$	$m/4$
B	0	0	$m/4$	$m/2$	m	$1-\frac{13}{4}m$	m	$m/2$
$CCC-C$	0	0	0	m	$2m$	$4m$	$1-15m$	$8m$
D/SD	0	0	0	0	0	0	0	1

The entry in the ith row of the jth column is the probability to migrate from class i to j over the next month. m is a parameter that refers to the basic migration rate into the neighboring classes and has to be specified. We choose $m = 0.003$ for sovereigns. Based on the empirical finding that firms have higher migration rates (resulting in higher PDs) we simply rescale the migration matrix for corporates by multiplying m with some constant k. In our simulations, we will consider $k = 1, 1.25, 1.5$. While the migration rates for the upper six categories follow the same pattern it was necessary to introduce higher migration rates for the CCC-C category to mirror the high CCC-C default rates observed empirically.

At first sight it might seem more appropriate to simply choose a migration matrix based on historical migration rates for the data-generating process. However, we have chosen not to do so because of two reasons. First, for reasons which we have discussed in the introduction, the implied pseudo-true PDs would be at an unrealistically low level. For instance, in this case we would have implied pseudo-true PDs of 0.06% for BB rated sovereigns at a one-year horizon and 0.13% for BBB rated sovereigns at a five-year horizon. These PDs are considerably lower than our standard estimates which - as will be confirmed by our simulations - already have a tendency to underestimate true PDs in small samples. Second, we want to investigate different scenarios for the difference between sovereigns and firms (specified by different choices for k) which is more straightforward within our setting.

Sovereign and corporate default and migration rates are very likely to be affected by common shocks like, for instance, recessions. We account for this kind of dependence by applying a CreditMetrics^{TM}-type approach (Gupton et al., 1997). The procedure involves simulating observations from a multivariate normal distribution and mapping these realizations to rating changes. Consider for example a AAA rated sovereign which has a probability to remain AAA over the next month of $1 - \frac{7}{4}0.003 = .99475$ and a probability to migrate to AA of 0.003. If the corresponding realization of the normal distribution is smaller than $\Phi^{-1}(.99475) \approx 2.5589$ the rating for the next month is set to AAA again. If

[15] These are calculated by exponentiating the migration matrix and are presented below together with the simulation results.

Table 4: Evaluation of the standard estimator

	Panel A: Pseudo-true PDs (%)				Panel B: Relative bias			
	1 year	3 years	5 years	10 years	1 year	3 years	5 years	10 years
AAA	4.7e-4	0.014	0.061	0.414	-0.213	0.044	0.083	0.058
AA	0.005	0.067	0.223	1.069	0.085	0.059	0.032	0.018
A	0.018	0.203	0.609	2.414	0.070	0.022	0.011	0.003
BBB	0.132	0.981	2.245	6.101	0.005	0.016	0.007	-0.001
BB	1.082	3.905	6.958	14.124	0.001	-0.005	-0.006	-0.008
B	2.122	7.355	12.615	23.544	0.007	0.004	0.001	-0.010
CCC-C	22.786	44.379	52.797	60.227	0.038	0.014	-0.004	-0.019

	Panel C: Relative RMSE				Panel D: % $\widehat{PD} < PD$			
	1 year	3 years	5 years	10 years	1 year	3 years	5 years	10 years
AAA	15.393	5.722	3.680	2.105	99.62	94.98	88.48	74.02
AA	7.271	3.529	2.525	1.693	97.06	87.98	80.08	68.32
A	3.685	1.871	1.440	1.123	89.68	71.64	63.90	59.30
BBB	1.358	0.914	0.801	0.719	61.74	56.62	56.56	56.50
BB	0.561	0.498	0.479	0.478	53.52	53.74	53.70	54.30
B	0.453	0.415	0.407	0.428	53.22	52.44	53.14	52.72
CCC-C	0.378	0.355	0.352	0.356	49.50	50.24	50.54	51.52

Relative bias and Relative Root Mean Squared Error (RMSE) are calculated as $\left(\widehat{PD} - PD\right)/PD$ and $RMSE/PD$, respectively. % $\widehat{PD} < PD$ is the percentage of simulations for which the estimated PD was below the pseudo-true PD. The number of simulations is 5000.

instead the realization is in the interval $[\Phi^{-1}(.99475), \Phi^{-1}(.99775)] \approx [2.5589, 2.8408]$ the sovereign migrates to AA, and so on. The correlations of the corresponding multivariate normal distribution have the same meaning as the so-called asset correlations that are part of the IRB formula in the Basel II framework.[16] There, the asset correlations are specified as a function of the one-year PD (BCBS, 2006a, §272). We adopt this approach to specifiy the correlations of our multivariate normal distribution.

For the sake of illustration, we will from now on concentrate on prediction horizons of 1, 3, 5 and 10 years. All upcoming results are based on 5000 simulations. We start the presentation of our simulation results with the evaluation of the standard estimator which is given in Table 4. Panel A of Table 4 shows the pseudo-true PDs implied by our data-generating process which are of similar magnitude as our empirical estimates but, importantly, are not zero even for the highest rating grades. In Panel B we see the estimated bias of the standard estimator relative to the pseudo-true values. While we know that the standard estimator is consistent it is interesting to see that there is also

[16]In both cases, they can be thought of as the correlations between the asset values of firms. Once these asset values cross a certain lower threshold, a default event occurs.

no significant bias (with the exception of the CCC-C category) in small samples.[17]

The accuracy of the standard estimator as measured by the Root Mean Squared Error (RMSE) relative to the pseudo-true PDs is shown in Panel C of Table 4. It is clearly visible that the estimation uncertainty rises in relative terms as the pseudo-true values decline. Therefore, especially in these cases there should be potential to improve upon the standard estimator. Finally, in Panel D we report the proportions of the simulations where the pseudo-true PD has been underestimated by the standard estimator. The fact that we observe values well above 50% is caused by the highly skewed sampling distribution of the standard estimator in small samples and especially under small true PDs. Note that this feature can also be interpreted as a kind of bias called median bias in the statistics literature. Following Birnbaum (1963), the median bias of a PD estimator is given as $P(\widehat{PD} > PD|PD) - P(\widehat{PD} < PD|PD)$, and the estimator is called median-unbiased if the median bias is equal to zero. Under this concept, although being mean-unbiased, the standard estimator is clearly downward median-biased which is an obvious drawback at least if conservativeness is among the criteria to evaluate an estimator.

We now turn to the evaluation of the EB estimator. Table 5 shows its precision as measured by the ratio of the RMSEs of the EB and the standard estimator so that values smaller than 1 indicate a superior performance of the EB estimator. We report results for two scenarios, $k = 1.25$ in Panel A and $k = 1.5$ in Panel B. To provide insight into the relative level of PDs implied by these specifications we show in the left parts of Table 5 the ratio of the pseudo-true PDs of corporates and sovereigns. For instance, the BBB one-year PD is 52% higher for corporates than for sovereigns under $k = 1.25$. Still, in this case the RMSE of the EB estimator is 44.5% lower than the RMSE of the standard estimator. Overall, we observe an improvement by using the EB estimator in nearly all cases with a few exceptions at $k = 1.5$. The relative strength of the EB estimator increases i) as the sample size decreases (as can be seen by the large improvements for the CCC-C category), ii) as the pseudo-true PDs decrease (see the robust EB performance with respect to the AAA and AA one-year PDs) and iii) as the distance between the corporate and the sovereign pseudo-true PDs decreases (Panel A vs. Panel B). Case i) is expected from theory and further confirmed by additional simulations which we do not report here but are available on request. In these simulations, we randomly dropped half of our sample from the simulations which is still likely to be a practically realistic sample size. Under this reduced sample the relative EB performance is even better. For instance, under $k = 1.5$, the RMSE ratio of the A one-year PD then decreases from 0.696 to 0.520. With respect to iii), it is important to note that beyond some point the EB estimator begins to get worse than the standard estimator. For example, when we set $k = 1.75$, the RMSE ratio is not in all but in most cases larger than 1.

[17] We explored the significance of the bias by using Monte Carlo standard errors. At a confidence level of $\gamma = 0.05$, the bias was only significant for the CCC-C PDs at horizons of 1, 3 and 10 years. The special role of the CCC-C category is not too surprising given that it has the smallest sample size.

Table 5: Precision of the empirical Bayes estimator

Panel A: $k = 1.25$

	Pseudo-true PD ratio				RMSE ratio			
	1 year	3 years	5 years	10 years	1 year	3 years	5 years	10 years
AAA	1.94	1.90	1.86	1.79	0.841	0.737	0.708	0.735
AA	1.65	1.68	1.67	1.61	0.814	0.656	0.608	0.642
A	1.61	1.62	1.58	1.50	0.650	0.584	0.621	0.648
BBB	1.52	1.45	1.40	1.34	0.555	0.597	0.611	0.605
BB	1.29	1.29	1.27	1.23	0.627	0.672	0.662	0.616
B	1.29	1.27	1.24	1.18	0.688	0.705	0.666	0.560
CCC-C	1.18	1.09	1.05	1.03	0.559	0.480	0.487	0.580

Panel B: $k = 1.5$

	Pseudo-true PD ratio				RMSE ratio			
	1 year	3 years	5 years	10 years	1 year	3 years	5 years	10 years
AAA	3.32	3.19	3.07	2.83	0.848	0.819	0.881	1.056
AA	2.49	2.57	2.52	2.34	0.830	0.746	0.819	0.955
A	2.39	2.38	2.27	2.04	0.696	0.884	0.965	0.980
BBB	2.13	1.94	1.83	1.67	0.826	0.930	0.946	0.883
BB	1.60	1.59	1.53	1.43	0.944	1.005	0.964	0.845
B	1.58	1.54	1.47	1.34	1.082	1.081	0.975	0.732
CCC-C	1.34	1.16	1.09	1.06	0.798	0.543	0.509	0.574

The pseudo-true PD ratio is calculated as $PD(\text{corporate})/PD(\text{sovereign})$ and the RMSE ratio is defined as $RMSE(\text{EB})/RMSE(\text{standard estimator})$. The number of simulations is 5000.

Table 6: Conservativeness of the empirical Bayes estimator

	% $\widehat{PD} < PD$ for $k = 1$				% $\widehat{PD} < PD$ for $k = 1.25$			
	1 year	3 years	5 years	10 years	1 year	3 years	5 years	10 years
AAA	95.6	73.5	60.4	52.5	91.2	54.9	38.3	28.7
AA	64.7	53.9	49.3	46.1	50.2	27.6	20.2	16.1
A	53.2	46.3	47.3	46.9	27.3	13.3	11.6	13.0
BBB	47.3	51.5	53.6	53.2	10.0	12.9	19.0	21.9
BB	53.3	53.7	53.7	53.6	19.5	19.5	21.4	25.7
B	54.3	54.8	53.8	53.7	16.5	16.8	18.8	24.4
CCC-C	53.2	50.7	46.9	40.6	12.8	23.3	30.5	31.3

% $\widehat{PD} < PD$ is again the percentage of simulations for which the estimated PD was below the pseudo-true PD. The number of simulations is 5000.

Depending on the true data generating process, the EB estimator may or may not be more precise than the standard estimator. More clearly, we can ascribe the EB estimator to be more conservative which can be seen from the results in Table 6. To evaluate the conservativeness of the EB estimator, we have chosen the scenarios $k = 1$ and $k = 1.25$. For larger values of k the conservativeness of the EB estimator will obviously further rise. But even for $k = 1$, the relative frequency of underestimating the pseudo-true PD is distinctively lower than for the standard estimator which was given in Table 4. Since no bias is introduced in the case of $k = 1$ the reason for this finding is just the less skewed sampling distribution of the EB estimator, an effect similar to the effect of an increasing sample size. In the case of $k = 1.25$, an additional upward bias is present so that underestimation of the pseudo-true PD happens only in less than 50% of all cases with a few exceptions for very small pseudo-true PDs.

The analysis of our estimators on a portfolio basis provides further insights. We again stick to our approximate market portfolio (see section 4) and consider the estimation of expected losses and capital requirements again assuming a recovery rate of 0.55. As was briefly mentioned in the introduction, the estimation of expected losses over the whole life of the portfolio to calculate loan loss reserves may gain importance due to recent regulatory efforts (BCBS, 2009). Panel A of Table 7 shows the performance of our estimators in predicting expected losses over various time horizons. Differently to the expected return calculations in section 4, since we now do not refer to any specific bond, we do not consider any coupon payments instead assuming only one hypothetical cash flow at the end of the prediction horizon. Interestingly, the standard estimator now improves relative to the EB estimator. The RMSE of the EB estimator is now lower than that of the standard estimator only for the scenario with $k = 1.25$, whereas the RMSE is now higher for $k = 1.5$. This is because, when estimating expected losses for a portfolio, there is a kind of variance reducing effect as compared to estimating expected losses for a single obligor.[18] No such reduction effect holds for the bias of the estimators. Since the EB estimators benefits hinge on its ability to reduce variance at the cost of some bias, the EB estimators merits diminish somewhat in this case. As far as conservativeness is concerned, the standard estimator is still liberal underestimating the pseudo-true expected losses in more than 50% of all simulations. In contrast, the EB estimator tends to overestimate pseudo-true expected losses.

Similarly to section 4, we also analyze our estimators with respect to implied Basel II capital requirements. More precisely, we investigate how good our estimators perform in estimating pseudo-true economic capital which we define as the capital requirements which follow from our pseudo-true PDs and, again, the IRB formula. The results of Panel B of Table 7 are astonishing. Now, the standard estimator has a large downward bias as pseudo-true economic capital is underestimated by 45% on average.[19] This finding

[18]Note that the standard error of the weighted average of some estimators is lower than the weighted average of their standard errors. This effect works here since the estimated portfolio-wide expected losses are weighted averages of the estimated rating-specific losses.

[19]The bias of the standard estimator in our expected loss calculations is, in contrast, negligibly small.

Table 7: Portfolio evaluation of estimators

Panel A: Estimation of expected losses for market portfolio

		Relative RMSE			$\% \widehat{EL} < EL$		
	Pseudo-true EL (%)	SE	EB (k=1.25)	EB (k=1.5)	SE	EB (k=1.25)	EB (k=1.5)
1 year	0.06	0.49	0.45	0.66	55.9	12.5	4.2
3 years	0.26	0.54	0.52	0.78	58.6	10.7	2.0
5 years	0.57	0.57	0.54	0.84	59.3	11.1	1.7
10 years	1.69	0.60	0.59	0.92	57.6	12.4	2.3

Panel B: Estimation of capital requirements for market portfolio

	Relative Bias			Relative RMSE			$\% \widehat{EC} < EC$		
Pseudo-true EC (%)	SE	EB (k=1.25)	EB (k=1.5)	SE	EB (k=1.25)	EB (k=1.5)	SE	EB (k=1.25)	EB (k=1.5)
0.79	-0.45	-0.01	0.16	0.57	0.32	0.36	94.6	57.8	32.4

The composition of the market portfolio is given in Table 3. Relative RMSE is defined as RMSE/Pseudo-true EL and RMSE/Pseudo-true EC, respectively, where EL means expected losses and EC means economic capital. $\% \widehat{EL} < EL$ and $\% \widehat{EC} < EC$ are the relative frequencies of simulations where pseudo-true EL/EC was underestimated. Relative bias is the estimated bias divided by pseudo-true EC. SE is the abbreviation for the standard estimator whereas EB refers to the empirical Bayes estimator. The number of simulations is 5000.

can be explained by the concavity of the IRB formula. Denote by $K(\cdot)$ the function that calculates economic capital using the PD as an argument. Then, by a simple second-order Taylor series expansion:

$$\begin{aligned} E[K(\widehat{PD}) - K(PD)] &\approx K'(PD) \cdot E[\widehat{PD} - PD] + \frac{1}{2} K''(PD) \cdot E[(\widehat{PD} - PD)^2] \\ &\approx \frac{1}{2} K''(PD) \cdot V[\widehat{PD}] < 0 \end{aligned}$$

The equation holds for an approximately unbiased estimator. Since the IRB function is concave we have $K''(PD) < 0$ which results in a downward biased estimate for economic capital under an unbiased estimate for the PD.[20] The bias is proportional to the variance of the estimator, $V[\widehat{PD}]$, and may thus only be negligible if estimation uncertainty is low. However, the results of Table 7 show that the opposite case is true and that the bias is substantial under a realistic scenario. In contrast, the upward bias of the EB estimator now compensates this effect and leads to nearly unbiased estimation for $k = 1.25$ and only slightly upward biased estimation for $k = 1.5$. Further, the precision of the EB

[20]This type of bias has already been explored by Kiefer & Larson (2003). In that study, the authors also calculate the second derivative of the IRB economic capital function.

estimator is now higher than for the standard estimator, even for $k = 1.5$. On top of that, the standard estimator now underestimates pseudo-true economic capital at an extremely high rate of 94.6% whereas the same figures are 57.8% and 32.4% for the EB estimator. Evidently, the EB estimator is more appropriate for economic capital calculations in our scenario.

6. Conclusions

The EB methods proposed in this paper offer an attractive combination of possible gains in precision and a pronounced increase in conservativeness as compared to standard approaches. This is remarkable since more conservative methods (like estimators based on upper confidence bounds) are usually connected with distinctively less precision when used as a point estimator. The crucial advantage of the EB estimator is that it adds valuable information from similar datasets to the analysis. Our simulation results back the hypothesis that the improvements of the EB estimator can be substantial and economically meaningful. Moreover, there are likely to be practical situations where the relative performance of the EB estimator is even better than in the cases studied by us. For instance, a larger number of different datasets, i.e. $G > 2$, will usually make the estimation of the prior parameters and thus the EB estimator more efficient. As we have already mentioned before, smaller sample sizes will also make the EB estimator more valuable in relative terms.

With respect to our application to sovereign bonds, we conclude that the zero default rates usually reported for AAA-A rated sovereigns should be interpreted with care. As our simulation study shows it is quite likely to observe no default event if the corresponding true PDs are small but posititive. More generally, our simulation results confirm the hypothesis that the standard estimator is anticonservative for sovereign PD estimation thereby failing to fulfill the Basel II requirements mentioned in section 1. Among our applications to credit portfolio risk, the results regarding the calculation of capital requirements are most notably. In section 4 we showed that small increases in (sovereign) PD estimates (by using the EB estimator) lead to large rise of implied capital requirements well above the level calculated under standard approaches. Then, in section 5, we have seen that the EB estimator leading to these increasing capital requirements is likely to be a better estimator for this purpose.

A R code for empirical Bayes estimator

The following code refers to Formulas (7)-(10):

```
  eb <- function(w,lambda,n,iter)  {
   # w: Vector of weights; sum(w) should be 1
   # lambda: Vector of conventional hazard rates (or PDs)
   # n: Vector of numbers at risk/sample sizes
   # iter: TRUE or FALSE, referring to iteration step
   G <- length(w)
   # G is the number of groups/portfolios
   mu <- sum(w*lambda)
   SS <- (G-1)/G*sum(w*(lambda-mu)^2)
   tau <- (SS-mu*(1-mu)*sum(w*(1-w)/n))/(mu*(1-mu)*sum((1-1/n)*w*(1-w)))
   tau <- ifelse(tau < 0,0,ifelse(tau > 1,1,tau))
   if (iter==TRUE) {
     w <- n/(1+tau*(n-1))/sum(n/(1+tau*(n-1)))
     mu <- sum(w*lambda)
     SS <- (G-1)/G*sum(w*(lambda-mu)^2)
     tau <- (SS-mu*(1-mu)*sum(w*(1-w)/n))/(mu*(1-mu)*sum((1-1/n)*w*(1-w)))
     tau <- ifelse(tau < 0,0,ifelse(tau > 1,1,tau))
   }
   B <- (1-tau)/(1+tau*(n-1))
   B <- ifelse(B<0,0,ifelse(B>1,1,B))
   lambda.eb <- B*mu + (1-B)*lambda
   if (sum(lambda)==0) lambda.eb <- lambda
   return(lambda.eb)
 }

 # Example:
 eb(c(1/2,1/2),c(0.04,0),c(1000,100),iter=TRUE)
  0.03875106 0.01130409
```

References

Altman, E. I. (1989). Measuring corporate bond mortality and performance. *Journal of Finance*, 44(4), 909–922.

BBA (British Bankers' Association, London Investment Banking Association, International Swaps and Derivatives Association) (2005). *Low Default Portfolios.* Joint Industry Working Group Discussion Paper (January 2005). Available at http://www.isda.org.

BCBS (Basel Committee on Banking Supervision) (2005). *Validation of Low-Default Portfolios in the Basel II Framework.* Basel Committee Newsletter No. 6 (September 2005).

BCBS (Basel Committee on Banking Supervision) (2006a). *International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version.* http://www.bis.org/publ/bcbs128.pdf.

BCBS (Basel Committee on Banking Supervision) (2006b). *The IRB Use Test: Background and Implementation*. Basel Committee Newsletter No. 9 (September 2006).

BCBS (Basel Committee on Banking Supervision) (2009). *Guiding Principles for the Replacement of IAS 39*. Available at http://www.bis.org/publ/bcbs161.htm.

BCBS (Basel Committee on Banking Supervision) (2010). *Group of Governors and Heads of Supervision Announces Higher Global Minimum Capital Standards*. Press release, September 12, 2010.

Benjamin, N., Cathcart, A., & Ryan, K. (2006). *Low Default Portfolios: A Proposal for Conservative Estimation of Default Probabilities*. Financial Services Authority, London. Discussion Paper.

Birnbaum, A. (1963). Median-unbiased estimators. *Bulletin of Mathematical Statistics*, 11, 25–34.

Carlin, B. P. & Louis, T. A. (2008). *Bayesian Methods for Data Analysis*. Boca Raton: CRC Press, 3rd edition.

Casella, G. (1985). An introduction to empirical Bayes analysis. *The American Statistician*, 39(2), 83–87.

Christensen, J. H., Hansen, E., & Lando, D. (2004). Confidence sets for continous-time rating transition probabilities. *Journal of Banking & Finance*, 28, 2575–2602.

Cutler, S. J. & Ederer, F. (1958). Maximum utilization of the life table method in analyzing survival. *Journal of Chronic Diseases*, 8(6), 699–712.

Fuertes, A.-M. & Kalotychou, E. (2007). On sovereign credit migration: A study of alternative estimators and rating dynamics. *Computational Statistics & Data Analysis*, 51, 3448–3469.

Gupton, G. M., Finger, C. M., & Bhatia, M. (1997). *CreditMetrics^{TM} - Technical Document*. J.P. Morgan, New York.

Hamilton, D. T. & Cantor, R. (2006). *Measuring Corporate Default Rates*. Moody's Investors Service. Special Comment.

Hanson, S. & Schuermann, T. (2006). Confidence intervals for probabilities of default. *Journal of Banking & Finance*, 30, 2281–2301.

Hjort, N. L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Annals of Statistics*, 18(3), 1259–1294.

Johansen, S. (1978). The Product Limit estimator as Maximum Likelihood estimator. *Scandinavian Journal of Statistics*, 5(4), 195–199.

Kiefer, N. M. (2009). Default estimation for low-default portfolios. *Journal of Empirical Finance*, 16(1), 164–173.

Kiefer, N. M. (2010). Default estimation and expert information. *Journal of Business & Economic Statistics*, 28(2), 320–328.

Kiefer, N. M. & Larson, C. E. (2003). *Biases in Default Estimation and Capital Allocations under Basel II.* Working paper.

Kleinman, J. C. (1973). Proportions with extraneous variance: Single and independent sample. *Journal of the American Statistical Association*, 68(341), 46–54.

Kurbat, M. & Korablev, I. (2002). *Methodology for Testing the Level of the EDF^{TM} Credit Measure.* Moody's KMV Technical Report No. 020729.

Lando, D. & Skodeberg, T. M. (2002). Analyzing Rating Transitions and Rating Drift with Continous Observations. *Journal of Banking & Finance*, 26, 423–444.

Lawrenz, J. (2008). Assessing the estimation uncertainty of default probabilities. *Kredit und Kapital*, 41(2), 217–238.

Pluto, K. & Tasche, D. (2006). Estimating probabilities of default for low default portfolios. In B. Engelmann & R. Rauhmeier (Eds.), *The Basel II Risk Parameters* (pp. 79–103). Berlin: Springer.

Standard & Poor's (2011a). *Sovereign Defaults and Rating Transition Data, 2010 Update.* Publication date: February 23, 2011.

Standard & Poor's (2011b). *Sovereign Rating and Country T&C Assessment Histories.* Publication date: June 3, 2011.

Standard & Poor's (2011c). *Standard & Poor's Noninvestment Grade Sovereign Debt Recovery Ratings: 2011 Update.* RatingsDirect, Global Credit Portal: March 29, 2011.

Tasche, D. (2011). *Bayesian Estimation of Probabilities of Default for Low Default Portfolios.* Available at arxiv.org.