Accepted Manuscript

Maximum likelihood estimation of first-passage structural credit risk models correcting for the survivorship bias

Diego Amaya, Mathieu Boudreault, Don L. McLeish

Please cite this article as: Diego Amaya, Mathieu Boudreault, Don L. McLeish, Maximum likelihood estimation of first-passage structural credit risk models correcting for the survivorship bias, *Journal of Economic Dynamics & Control* (2019), doi: https://doi.org/10.1016/j.jedc.2018.11.005

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Maximum likelihood estimation of first-passage structural credit risk models correcting for the survivorship bias

Diego Amaya

Department of Finance
Lazaridis School of Business & Economics
Wilfrid Laurier University
Waterloo, ON, Canada
Email: damaya@wlu.ca

Mathieu Boudreault [*]

Department of Mathematics
Faculty of Science
Université du Québec à Montréal
Montréal, QC, Canada
Email: boudreault.mathieu@uqam.ca

Don L. McLeish

Department of Statistics and Actuarial Science
Faculty of Mathematics
University of Waterloo
Waterloo, ON, Canada
Email: dlmcleis@math.uwaterloo.ca

September 1st, 2018

1

Abstract

The survivorship bias in credit risk modeling is the bias that results in parameter estimates when the survival of a company is ignored. We study the statistical properties of the maximum likelihood estimator (MLE) accounting for survivorship bias for models based on the first-passage of the geometric Brownian motion. We find that if we neglect the survivorship bias, then the drift has a positive bias that may not disappear asymptotically. We show that correcting the survivorship bias by conditioning on survival in the likelihood function underestimates the drift. Therefore, we propose a bias correction method for non-iid samples that is first-order unbiased and second-order efficient. The economic impact of neglecting or miscorrecting for the survivorship bias is studied empirically based on a sample of more than 13,000 companies over the period 1980 through 2016 inclusive. Our results point to the important risk of misclassifying a company as solvent or insolvent due to biases in the estimates.

Keywords: survival bias, geometric Brownian motion, conditional estimation, default probability, inference, diffusion processes

JEL classification: C01, C13, C18, C51, G12.

2

1 Introduction

Survivorship bias (or survival bias[1]) is the error that emerges when a study is solely based on a sample of observations that exclude failures. This bias is well documented in finance (e.g. Brown et al. (1995)), notably in performance studies (e.g. Brown et al. (1992) and Elton et al. (1996)) i.e. companies that go out of business are usually removed from the sample.

Quantifying a firm's solvency is a very important task in risk management. However, credit risk analysis and estimation could also be subject to survivorship bias because ignoring a company's survival during the observation period may overestimate its solvency. The aim of this paper is to analyze maximum likelihood estimation (MLE) of first-passage structural credit risk models accounting for survivorship. We show that maximum likelihood inference under survivorship is a non-trivial statistical problem and that there are cases in which the obvious estimator, the survivorship-corrected MLE, suffers from substantial biases. We propose a methodology to adjust for this bias, and conduct simulated and empirical studies about the economic impact of not accounting for these biases.

It is common practice to assess the credit risk of a company from its stock prices. Structural models are used for this purpose because they link credit risk to a company's capital structure (assets, liabilities, equity). The credit risk model proposed in Merton (1974) assumes that default only occurs at debt's maturity when assets are insufficient to repay debtholders. In this model, there is no survivorship bias because assets can take any value prior to debt's maturity. However, this default triggering mechanism is not realistic because creditors/investors will never let a debtor's solvency degrade to a point where they could lose invested capital significantly. First-passage structural credit risk models assume instead that default occurs the first time the asset value process crosses a deterministic barrier (see e.g. Black and Cox (1976)), thus allowing for default before maturity. In this context, a survivorship bias appears because observing the company's equity prices today is necessarily conditional upon its survival.

The literature in finance has largely focused on estimating the parameters of Merton

---

[1]We will use the terms "survivorship bias" and "survival bias" indistinguishably throughout this article.

(1974) because the capital structure is simple and yields closed-form expressions for equity prices. Examples of these works include Jones, Mason & Rosenfeld (1984), in which the authors set up a system of two equations and two unknowns based upon the initial equity value and stock volatility to back up the initial asset value and volatility. Although appealing due to its simplicity, Ericsson & Reneby (2004) have shown that it yields severe biases. Duan (1994, 2000) present an MLE-based estimation of the model in which the equity price is treated as a one-to-one transformation of the assets. Vassalou & Xing (2004) propose an alternative technique that uses an iterative algorithm to first find the asset volatility and subsequently the drift of the process – a procedure similar to the one employed in Moody's KMV model. As discussed in Duan et al. (2005), this technique is equivalent to the expectation-maximization (EM) algorithm, thus yielding estimates that are equivalent to the MLE of Duan (1994).

When estimating first-passage models, most authors have relied on the aforementioned techniques and have thus tended to neglect the survival bias (see e.g. Wong and Choi (2009), Dionne and Laajimi (2012), Afik et al. (2016)). There are, however, two noteworthy exceptions. In Duan et al. (2004), the authors explicitly account for the survivorship bias when estimating the parameters of a Merton-like model with multiple refinancing dates. Forte & Lovreta (2012) apply Duan et al. (2004) and investigate other approaches to infer the parameters of a first-passage model with credit default swap premiums. Unlike these authors, we focus on analyzing the statistical properties of the MLE for first-passage structural models conditional upon survival and present an extensive empirical study with a very large sample of companies to assess the impact of survivorship.

Our analysis is also related to the work of Li, Pearson and Poteshman (2004). In a class of diffusion processes, they derive a set of moments conditional on typical events (such as survival) for the purpose of GMM estimation. Whereas they briefly discuss the issue of survivorship bias, they do not formally analyze properties of the maximum likelihood estimators nor quantify this bias empirically.

The broader problem that we investigate is the MLE of a geometric Brownian motion (GBM) conditional upon its infimum never crossing a fixed known barrier. The first contribution of the paper is to analyze the impact of ignoring the survivorship bias on the MLE of

4

the drift and diffusion of this process. We show that the survivorship bias has a non trivial impact on the estimate of the drift. We compute explicitly the size of the bias and show that it is always positive (drift is overestimated). Also, we demonstrate that this bias only disappears asymptotically if the drift is positive.

Our second contribution is the characterization of the statistical behavior associated with the conditional MLE (i.e. conditional upon survival). We work with this likelihood function to show that the final asset value is a sufficient statistic to estimate the drift, just like in the unconditional case. We use this characterization to show that the conditional MLE is biased downward, so numerical maximization of the likelihood function yields biased estimates of the drift[2].

The third contribution is a method to correct for the survivorship bias when estimating the drift. Standard bias correction methods such as the jackknife or bootstrap do not work in this context because observations are not independent nor identically distributed (iid). Based upon a series expansion of the expected parameter estimate (see e.g. Cox & Hinkley (1974)), we propose a debiasing method that is first-order unbiased and second-order efficient, which can be applied to correct the survivorship bias. This method adds to a rich literature in financial econometrics that analyzes and corrects some biases in the parameters of financial models (see e.g. Tang and Chen (2009), Yu (2012) and Bauer et al. (2012)).

Our fourth contribution is an extensive simulated and empirical study whose goal is to analyze the economic impact of ignoring or miscorrecting for survivorship bias in credit risk analyses. Based on a sample of 13,794 firms obtained from the intersection of CRSP and Compustat, we use monthly stock prices between 1979 and 2016 to estimate a credit risk model for each firm in the sample. We compare default probabilities and the proportion of risky companies over different business cycles between 1980 and 2016. We find that biases in the naive (unconditional) and conditional MLEs are significant economically. For example, during the recession of the early 1980s, the one-year default probability of a company is 25% when estimated with a debiased MLE, which is considerably higher than the 15% value obtained with the unconditional MLE, but lower than the 40% value obtained

---

[2]This downward bias has also been observed numerically in Duan et al. (2005). The authors attribute this bias to the length of the time horizon but we are able to show that the downward bias remains, as long as the time horizon is finite.

with the conditional MLE. These values illustrate how unconditional estimates are overly optimistic about the solvency of a company and conditional MLEs generally provide the opposite picture. Moreover, these effects are more important during economic recessions. The findings of this paper provide new evidence of how the survivorship bias can undermine risk management policies.

The remainder of the article is organized as follows. Section 2 introduces the general modeling framework and highlights the difficulties that arise with the maximum likelihood inference of the geometric Brownian motion conditional upon survival. Then, in Section 3, we introduce a debiasing method and we analyze numerically the behavior of this debiased estimator. Section 4 presents our credit risk modeling framework and how estimation should be conducted in such setting. Section 5 is an empirical study that illustrates the economic impact of ignoring or miscorrecting for the survivorship bias. Finally, Section 6 concludes whereas appendices comprise details of the proofs and additional numerical analyses.

2 Problem statement

The objective of this section is to analyze the MLE of the GBM under survivorship and highlight the problems that appear in that context. We first begin by laying the grounds of our modeling framework. We then characterize the bias in the MLE and the conditional MLE upon survivorship.

2.1 Framework

Let $\underline{A} = \{A_t, t \geq 0\}$ be a stochastic process that represents the value of a company's assets over time. We assume that $\underline{A}$ is a geometric Brownian motion (GBM). Written as a stochastic differential equation, the dynamics of $\underline{A}$ is

$$dA_t = \mu A_t dt + \sigma A_t dW_t,$$

where μ and σ are the drift and diffusion coefficients of the process. Moreover, the initial asset value $A_0 := a_0$ is known and $\{W_t, t \geq 0\}$ is a standard Brownian motion. Therefore,

6

the solution of the process is such that $\ln(A_t)$ is a standard Brownian motion (with drift) starting from $\ln(a_0)$ i.e.

$$\ln(A_t) = \ln(a_0) + \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t.$$

The second component of the modeling framework is the default barrier $L < A_0$. This parameter represents the amount owed by the debtor or any quantity deemed to represent a solvency limit for the company. When the company's asset value falls below or reaches L, it is assumed that the company defaults and stops any economic activity. Moreover, the default barrier L is known in advance or has to be determined exogenously (not estimated from $\underline{A}$).

Throughout the paper, we will often equivalently consider the stochastic process $\underline{Z} = \{Z_t, t \geq 0\}$ where

$$Z_t = \ln(A_t/L).$$

Therefore, $\underline{Z}$ is a Brownian motion with drift $\nu = \mu - \frac{1}{2}\sigma^2$, diffusion σ, and initial value of

$$z_0 := Z_0 = \ln(A_0/L) > 0.$$

The process $\underline{Z}$ is similar to the distance-to-default metric popular in credit risk modeling.[3]

We are interested in studying the stochastic process $\underline{A}$ subject to never crossing a barrier L. Therefore, let I_T^A be the minimum value attained by $\underline{A}$ over the time interval $[0, T]$ i.e.

$$I_T^A = \inf\{A_t, 0 \leq t \leq T\}$$

or equivalently

$$I_T^Z = \inf\{Z_t, 0 \leq t \leq T\}.$$

We refer to the event $I_T^A > L$ (or $I_T^Z > 0$) as the survival event, which is a fundamental part of first-passage structural credit risk models (see Section 4). In this setting, it becomes clear

---

[3]The common distance-to-default metric is usually normalized by the drift and diffusion (see Vassalou and Xing, 2004).

that selecting a company at time T is subject to survival.

This paper is concerned with the maximum likelihood estimation of ν and σ given survival. Assume first that we can directly observe the conditional process[4]

$$\underline{Z}|I_T^Z > 0$$

at discrete time points $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$, where n is the number of observations in the time interval $[0, T]$. For simplicity, assume time intervals have the same length i.e. that $h := t_i - t_{i-1} = T/n, \forall i = 1, 2, \ldots, n$.

2.2 Naive MLE

We now study parameter estimates of the process $\underline{Z}$ that ignore any survivorship consideration. These estimates are referred to as the naive MLE (NMLE) and they are denoted by $\widehat{\nu}^N$ and $\widehat{\sigma}^N$. They are obtained from the realizations of

$$R_i = Z_{t_i} - Z_{t_{i-1}} = \ln(A_{t_i}/A_{t_{i-1}}),$$

which correspond to the log-return of asset values over the time interval $[t_{i-1}, t_i]$. Since R_i is normally distributed with mean νh and variance $\sigma^2 h$, the NMLE of ν and σ are calculated by directly maximizing the joint unconditional normal p.d.f. $f_{R_1, R_2, \ldots, R_n}(r_1, r_2, \ldots, r_n)$.

2.2.1 Drift coefficient

The NMLE of ν is given by

$$\widehat{\nu}^N = \frac{1}{T} \ln(A_T/A_0) = \frac{1}{T}(Z_T - Z_0). \tag{2.1}$$

Since this estimate does not account for the survivorship event, there is a bias inherited in its computation. The following proposition establishes this bias.

---

[4]Later in Section 4, we will relax this assumption and suppose that we only observe a set of stock prices, obtained as a one-to-one transformation of asset values.

	1-year horizon					10-year horizon				
$\mu \mid A_0$	110	150	200	250	300	110	150	200	250	300
-0.1	0.2739	0.0455	-0.0642	-0.0929	-0.0987	0.1089	0.0791	0.0528	0.0334	0.0183
-0.05	0.2926	0.0733	-0.0231	-0.0452	-0.0492	0.1213	0.0920	0.0668	0.0487	0.0349
0	0.3127	0.1032	0.0198	0.0031	0.0005	0.1370	0.1085	0.0848	0.0685	0.0565
0.05	0.3342	0.1352	0.0643	0.0520	0.0503	0.1567	0.1293	0.1078	0.0938	0.0841
0.1	0.3574	0.1694	0.1101	0.1012	0.1001	0.1812	0.1552	0.1365	0.1254	0.1182
0.15	0.3821	0.2057	0.1569	0.1507	0.1501	0.2108	0.1866	0.1712	0.1631	0.1585
0.2	0.4085	0.2441	0.2047	0.2004	0.2000	0.2454	0.2234	0.2114	0.2061	0.2035
0.25	0.4366	0.2843	0.2531	0.2503	0.2500	0.2843	0.2645	0.2557	0.2525	0.2512
0.3	0.4664	0.3263	0.3020	0.3001	0.3000	0.3265	0.3089	0.3027	0.3010	0.3004

Table 1: Conditional expected value of $\widehat{\mu}^N$ for various values of μ and A_0. One-year (ten-year) horizon shown in the left (right) panel ($T = 1, 10$). Other parameters: $\sigma = 0.3$ and $L = 100$.

Proposition 2.1. *The NMLE of ν has a positive bias on a finite horizon T. It is given by*

$$
\mathrm{E}\left[\widehat{\mu}^N \,\middle|\, I_T^A > L\right] - \mu = \mathrm{E}[\widehat{\nu}^N | I_T^Z > 0] - \nu = \frac{2z_0}{T \Pr(I_T^A > L)} \exp\left(\frac{-2z_0\nu}{\sigma^2}\right) \Phi\left(\frac{\nu T - z_0}{\sigma\sqrt{T}}\right),
$$

where the survival probability is

$$
\Pr(I_T^A > L) = \Pr(I_T^Z > 0) = \Phi\left(\frac{\nu T + z_0}{\sigma\sqrt{T}}\right) - \exp\left(\frac{-2z_0\nu}{\sigma^2}\right) \Phi\left(\frac{\nu T - z_0}{\sigma\sqrt{T}}\right) \tag{2.2}
$$

and Φ is the normal c.d.f. (Proof shown in Appendix B.1) □

Since $z_0 > 0$, Proposition 2.1 shows that ignoring survivorship in the computation of the MLE always overestimates the drift of $\underline{A}$. In the context of credit risk modeling, this bias has important repercussions since default probabilities (based on these estimates) would be underestimated. To illustrate the size of this bias numerically, Table 1 shows $\mathrm{E}\left[\widehat{\mu}^N \,\middle|\, I_T^A > L\right]$ (i.e. the conditional expected value of $\widehat{\mu}^N$) for various combinations of A_0, μ and T, providing different default probability levels. For instance, over a 1-year horizon, the default probability for the pair ($\mu = -0.1, A_0 = 110$) is 85.34% and it decreases as we go further to the right and/or bottom in each panel of the table. In the most solvent case, the default probability is close to 0%.

First observe how the survivorship bias translates into a systematic overestimation of the

9

drift. As the default probability increases (decreasing μ and/or A_0), so does the bias size. Over a short time horizon (columns 2 to 6), the survivorship bias is very small whenever A_0 is very large. This is the case since the survival probability is significantly high for all values of μ considered (e.g. with $A_0 = 300$ and $\mu = -0.1$ the default probability is 0.13%).

When the time horizon increases (columns 7 to 11), the default probability is driven by the drift of the process. For instance, with $\mu < \frac{1}{2}\sigma^2$, the GBM is pulled down over time no matter what is the initial asset value, thus the firm will default with certainty. This behavior is observed in the right panel of Table 1 as the bias on the drift increases even if A_0 is very large. The following proposition characterizes this case.

Proposition 2.2. *The NMLE of ν is not consistent. Indeed, when $T \to \infty$:*

1. *$\widehat{\nu}^N \to \nu$ if $\nu \geq 0$.*

2. *$\widehat{\nu}^N \to 0$ if $\nu < 0$.*

(Proof shown in Appendix B.2) $\square$

Proposition 2.2 shows that as $T \to \infty$, the survivorship bias will only disappear when $\mu \geq \frac{1}{2}\sigma^2$ ($\nu \geq 0$). Otherwise, the GBM is pulled down by the drift but the conditioning event (survivorship) pushes the process upward. These are two opposing forces and the process ultimately approaches a Brownian excursion.

We now illustrate the economic impact of ignoring the survivorship bias. According to a 2015 study by Standard & Poor's[5], the 1-year default probability of non investment-grade bonds is in the range of 1% to 30%. Based on values presented in Table 1, if we fix $A_0 = 150$, the true default probabilities are in the range of non investment-grade. Let us consider a fictitious company whose asset drift is $\mu = 0.05$ (or $\nu = 0.05 - \frac{1}{2}\sigma^2 = 0.005$). The true default probability for this firm will be 17%, which corresponds to a rating somewhere between B and CCC. From the table, we see that ignoring survivorship yields an expected drift μ of 0.1352, with an associated default probability of 11%. This is an important underestimation of the default probability. This example shows that the survivorship bias on the drift could

---

[5]See Table 21 of S&P (2016)

significantly impact the *perceived* credit quality of a company and eventually tarnish an investor's risk management policy.

Overall, ignoring survivorship can affect the assessment of a company's solvency in a non trivial way, which points out the need to account for this bias in the inference of first-passage structural models.

2.2.2 Diffusion coefficient

The NMLE of σ^2 corresponds to

$$\left(\widehat{\sigma}^N\right)^2 = \frac{1}{(n-1)h} \sum_{i=1}^{n}(r_i - \overline{r})^2 \tag{2.3}$$

with

$$\overline{r} = \frac{1}{n} \sum_{i=1}^{n} r_i = \frac{Z_T - z_0}{n}.$$

Directly assessing the bias on this parameter requires the joint distribution of the random variables $(R_1, R_2, ..., R_n)$ and the firm's survival over $[0, T]$. Since this quantity is difficult to derive analytically, we rely instead on other properties of this estimate to analyze the impact of the survivorship bias. As noted in Merton (1980), increasing the sampling frequency provides more reliable estimates of σ^2. This result hinges on the fact that the quadratic variation of the process converges almost surely, so this estimator is almost surely consistent

$$(\widehat{\sigma^2})^N \to \sigma^2 \ \text{ as } h \to 0$$

for small h. Thereby, as the observation frequency of the process increases, survivorship considerations become less important for the diffusion coefficient.

To verify this, we perform a simulation exercise in which we compute the NMLE of σ^2 for different sample frequencies and sizes. Table 2 shows $\text{E}\left[\left(\widehat{\sigma}^N\right)^2 \middle| I_T^Z > 0\right]$ for $\sigma^2 = 0.04$ and $\sigma^2 = 0.16$ at various sampling frequencies (monthly (12), weekly (50) and daily (250)) and over different time horizons (1, 5, 10, and 25 years). All computations are carried out with 10,000 simulations. The initial asset value is set equal to $A_0 = 110$ to provide high default probabilities and hence a greater potential for survivorship bias.

11

	$\sigma^2 = 0.04$				$\sigma^2 = 0.16$			
T	Monthly	Weekly	Daily	Def prob	Monthly	Weekly	Daily	Def prob
1	0.0381	0.0393	0.0399	55.71%	0.1525	0.1566	0.1593	81.65%
5	0.0394	0.0398	0.0400	72.06%	0.1578	0.1594	0.1599	92.05%
10	0.0396	0.0399	0.0400	75.55%	0.1594	0.1595	0.1599	94.54%
25	0.0399	0.0399	0.0400	77.95%	0.1593	0.1598	0.1599	96.75%

Table 2: Conditional expected value of $\left(\widehat{\sigma}^N\right)^2$ for $\sigma^2 = 0.04, 0.16$ at various sampling frequencies (monthly (12), weekly (50) and daily (250)) and over different time horizons (1, 5, 10, 25 years). Other parameters employed in the simulation are: $\mu = 0.07$, $A_0 = 110$, $L = 100$. Last column (Def. prob.) gives the probability that the asset crosses the barrier in the interval $[0, T]$.

Results in Table 2 show that even if we ignore survivorship when estimating σ, the bias decreases quickly with the sampling frequency. The simulation exercise has been repeated for different values of μ (positive or negative) and A_0 yielding similar results. Therefore, our analyses suggest that we can obtain nearly unbiased estimates of σ^2 under survivorship by using $\widehat{\sigma}^N$ together with high frequency sampling. Alternatively, the bias can be made very small with weekly or monthly sampling as long as T increases.

In summary, σ can still be estimated very accurately with a reasonable number of observations even while ignoring survivorship. Consequently, in this section and in the following, we will assume that σ is known and focus our attention on the MLE of ν which has been shown to be much more sensitive to survival. Later in Section 4, we relax this assumption and estimate both parameters.

2.3 Conditional MLE

Instead of maximizing the joint *un*conditional density function to obtain the NMLE, which is the source of the survivorship bias, we now condition on survival of the company. That is, the conditional MLE (upon survivorship) should be based upon the joint conditional p.d.f.

$$f_{Z_{t_1}, Z_{t_2}, \ldots, Z_{t_n} \mid I_T^Z > 0}(z_1, z_2, \ldots, z_n)$$

where z_i is an observation of the r.v. Z_{t_i} for $i = 1, 2, \ldots, n$. The conditional likelihood function is provided in the following proposition.

Proposition 2.3. *The conditional likelihood function of the parameter ν given that the company survived in the time interval $[0, T]$ is*

$$\mathcal{L}(\nu) := \frac{1}{\Pr(I_T^Z > 0)} \times \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{\Delta t}} \phi\left(\frac{z_i - z_{i-1} - \nu\Delta t}{\sigma\sqrt{\Delta t}}\right) \left(1 - \exp\left(\frac{-2z_i z_{i-1}}{\sigma^2 \Delta t}\right)\right) \quad (2.4)$$

and ϕ is the normal p.d.f. When $z_n > 0$, this function converges to 0 when $\nu \to \pm\infty$ and therefore ν has a finite MLE. (Proofs shown in Appendix C.1 and C.3) $\square$

We denote by $\widehat{\nu}$ the conditional MLE (CMLE) of ν obtained by maximizing the likelihood function given in Equation (2.4). Since the CMLE of ν exists, typical numerical methods can be used to find $\widehat{\nu}$. Nonetheless, to understand the behavior of this estimate we characterize it with the help of the following proposition.

Proposition 2.4. *The CMLE $\widehat{\nu}$ obtained by maximizing Equation (2.4) is exactly equivalent to solving for ν in the expression*

$$z_n = \mathrm{E}[Z_T | I_T^Z > 0]$$

where

$$\mathrm{E}[Z_T | I_T^Z > 0] = z_0 + \nu T + 2z_0 \left(\frac{\Phi\left(\frac{z_0 + \nu T}{\sigma\sqrt{T}}\right)}{\Pr(I_T^Z > 0)} - 1\right).$$

Moreover, Z_T is a complete sufficient statistic for the parameter ν. (Proof shown in Appendix C.2) $\square$

Proposition 2.4 shows that despite having knowledge of the entire path $\underline{Z}$, only the initial and final observations z_0 and z_n matter in determining the CMLE of ν, just like in the case of the naive estimator $\widehat{\nu}^N$. This results in the following undesired properties.

Proposition 2.5. *Further properties of the CMLE*

1. $\mathrm{E}[\widehat{\nu} | I_T^Z > 0] - \nu \leq 0$ i.e. the bias of $\widehat{\nu}$ is always negative.

13

	1-year horizon					10-year horizon				
$\mu \mid A_0$	110	150	200	250	300	110	150	200	250	300
-0.1	-0.4668	-0.3767	-0.2300	-0.1479	-0.1164	-0.2883	-0.2862	-0.2815	-0.2758	-0.2697
-0.05	-0.3845	-0.2942	-0.1565	-0.0866	-0.0619	-0.2012	-0.1988	-0.1936	-0.1873	-0.1807
0	-0.3035	-0.2136	-0.0861	-0.0275	-0.0085	-0.1167	-0.1142	-0.1086	-0.1019	-0.0950
0.05	-0.2239	-0.1351	-0.0187	0.0295	0.0440	-0.0360	-0.0334	-0.0276	-0.0209	-0.0142
0.1	-0.1458	-0.0588	0.0460	0.0849	0.0958	0.0400	0.0426	0.0482	0.0544	0.0603
0.15	-0.0693	0.0151	0.1080	0.1390	0.1470	0.1105	0.1129	0.1179	0.1231	0.1277
0.2	0.0055	0.0866	0.1678	0.1921	0.1980	0.1755	0.1775	0.1816	0.1855	0.1887
0.25	0.0786	0.1558	0.2256	0.2444	0.2486	0.2355	0.2370	0.2401	0.2428	0.2448
0.3	0.1498	0.2225	0.2817	0.2960	0.2991	0.2916	0.2928	0.2950	0.2967	0.2978

Table 3: Conditional expected value of $\widehat{\mu}$ for various values of μ and A_0. One-year (ten-year) horizon shown in the left (right) panel ($T = 1, 10$). Other parameters: $\sigma = 0.3$ and $L = 100$.

2. When $z_n \to 0$, then $\widehat{\nu} \to -\infty$.

3. Moreover,

$$\frac{\mathrm{E}[\widehat{\nu}\,\mathbb{I}_{Z_T<\delta}]}{\nu} \to \infty$$

and hence

$$\mathrm{E}[\widehat{\nu}\,\mathbb{I}_{Z_T<\delta}] - \nu \to -\infty.$$

This essentially states that when $\nu \to -\infty$, then the bias in $\widehat{\nu}$ goes to $-\infty$.

(Proofs shown in Appendix C.4) □

A consequence of the first part of Proposition 2.5 is that correcting for the upward bias in the drift due to survivorship introduces a negative bias in the CMLE. The second and third statement in this proposition provide an explanation about the origin of this negative bias: if a path of a GBM approaches the barrier toward the end of the time interval, then even if ν is reasonably high, $\widehat{\nu}$ will tend to $-\infty$. Hence, whenever the likelihood of such event is significant, the CMLE will likely introduce a large negative bias.

Table 3 quantifies this negative bias by showing the conditional expected value of $\widehat{\mu}$ for the same combinations of μ and A_0 as in Table 1[6]. As it was the case with the NMLE, the default probability is also an important driver of the bias in the CMLE. Over a one-year

---

[6]Calculations in Table 3 are carried out with numerical integration as explained in computations of Equation (3.1).

horizon, we see that the CMLE has an important negative bias that disappears with the solvency of the firm i.e. when we increase μ and/or A_0. When the time horizon is $T = 10$, the default probability increases and so does the number of paths approaching the barrier at T from above. The results of Table 3 illustrate that even in the case of a larger time horizon, the bias still remains significant.

We now assess the economic impact of using the CMLE to correct for the survivorship bias within the context of a fictitious firm. For illustrative purposes, we set again the initial asset value to $A_0 = 150$ and the drift to $\mu = 0.05$. Table 3 shows that the expected CMLE in this case is -0.1351 (recall that the expected NMLE was $+0.1352$). With $\mu = -0.1351$, the resulting default probability is 36%, which is more than twice the real default probability of 17%. In this case, using the CMLE grossly overestimates the default probability of the company.

3 Bias correction method

We have shown in Section 2 that when we ignore survivorship (NMLE), the drift is overestimated but if we condition on the event that the barrier has never been crossed (CMLE), the drift is very much underestimated. Hence, we need to remove the bias in the CMLE.

Typical bias correction methods such as the bootstrap or jackknife resampling are not applicable in our context because observations from the conditional stochastic process $\underline{A}|I_T^A > L$ are not iid. Therefore, we propose a method for debiasing a MLE in non-iid samples based on a series expansion (see Cox and Hinkley (1974), Section 8.4). This bias correction method is presented in the next proposition and applied to the CMLE in the next subsection.

Proposition 3.1. *Suppose that $\widehat{\theta}$ is a biased MLE so that we can write*

$$\mathrm{E}[\widehat{\theta}] = g(\theta) = \theta + \frac{b_1(\theta)}{n} + \frac{b_2(\theta)}{n^2} + \ldots$$

Then the estimator $\widetilde{\theta}$ of θ obtained as

$$\widetilde{\theta} = g^{-1}(\widehat{\theta})$$

15

is first-order unbiased and second-order efficient. (Proof shown in Appendix D) □

Proposition 3.1 is general and applicable to other problems as well. It can also be used to remove the bias on the estimator of a one-to-one transformation $p(\theta)$ in a similar fashion.

When needed, inference on $\widetilde{\theta}$ is achieved by viewing this estimator as a one-to-one transformation of the MLE $\widehat{\theta}$. Therefore, a typical delta method can be used for example to obtain a confidence interval on $\widetilde{\theta}$.

3.1 Application to the conditional MLE

We now remove the bias in $\widehat{\nu}$ by applying Proposition 3.1 to the CMLE. Let us first define

$$\mathcal{E}(\nu) := \mathrm{E}[Z_T | I_T^Z > 0].$$

The key quantity that we need to compute in the debiasing method is $\mathrm{E}[\widehat{\nu}]$. But a difficulty that arises is the fact that for some realization z_n of Z_T (conditional upon survival), the CMLE

$$\widehat{\nu} = \mathcal{E}^{-1}(z_n)$$

has no analytical expression (see propositions 2.4 and 2.5 as well). Therefore,

$$\mathrm{E}[\widehat{\nu}] = \mathrm{E}[\mathcal{E}^{-1}(Z_T) | I_T^Z > 0] = g(\nu) \tag{3.1}$$

requires the numerical integration of an integrand whose points require numerical inversion. Moreover, $\widetilde{\nu}$, defined as the debiased CMLE of ν, requires the inversion of the function $g(\nu)$ which shows that debiasing is a computationally intense procedure. In the paragraphs that follow, we provide algorithms for the numerical implementation of the debiasing method whereas in Appendix D.2 we provide additional details and illustrations.

Algorithm 1: Computing $g(\nu)$ for one single ν

1. Compute $\mathcal{E}(\nu)$ for a fine and finite grid over $\nu \in [\underline{\nu}, \overline{\nu}]$. Since there is a closed-form expression for $\mathcal{E}(\nu)$, this step is done quickly.

16

2. Numerical integration:

   (a) Discretize the domain of the random variable $Z_T | I_T^Z > 0$ over $[0, \overline{z}]$ using a grid of (equally-spaced) points[7].

   (b) For each point $z \in [0, \overline{z}]$:

       i. Compute the corresponding CMLE $\mathcal{E}^{-1}(z)$ using interpolation from the grid in step 1.

       ii. Compute the integrand as the product of the CMLE (from step i.) times the conditional (upon survivorship) probability density function of Z_T (valued at z)

   (c) Then, perform the numerical integration using Newton-Cotes formulas and the previously computed integrand over $[0, \overline{z}]$.

To visualize the algorithm, Figure 1 shows the behavior of $g(\nu)$ for three values of z_0 whereas the thin dotted line is a 45 degree line. When the true value of ν is large, $g(\nu) \approx \nu$ which indicates the CMLE is almost unbiased. As we decrease ν, we see the extent of the bias as compared to the 45 degree line (no bias).

The last step of the algorithm is equivalent to finding the value of ν on the x-axis such that $g(\nu) = \widehat{\nu}$ on the y-axis. For instance, suppose we found in our sample that $\widehat{\nu} = 0$. Then we know it would be biased downward and that the size of the bias would be proportional to the default probability. When $A_0 = 300$, then the debiasing is minimal. However, if $A_0 = 110$, then ν is closer to 0.2 (rather than 0) to correct for a greater bias in $\widehat{\nu}$. Therefore, the debiased CMLE of ν, $\widetilde{\nu}$, is obtained with the following algorithm.

Algorithm 2: Debiaising the CMLE

1. For a given dataset, compute the CMLE $\widehat{\nu}_{obs}$.

2. With a numerical root-finding algorithm, find ν until $\widehat{\nu}_{obs} = g(\nu)$ using the algorithm for $g(\nu)$.

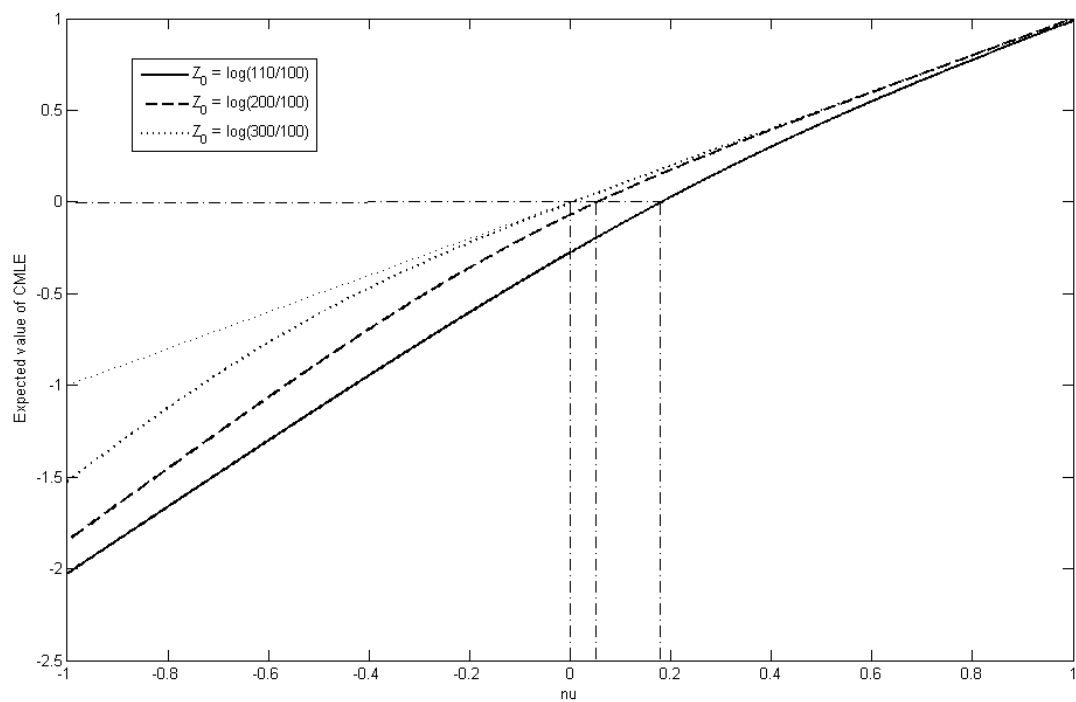[7]How this partition is built ultimately depends on the quadrature method used in the sub-step (c).

Figure 1: Expected CMLE as a function of ν for three values of z_0. Other values: $\sigma = 0.3, L = 100, T = 1$

18

3. Set $\widetilde{\nu}_{obs}$ as the value of ν found in step 2. This is the debiased CMLE of ν in the dataset.

3.2 Behavior of the estimator

We conduct numerical tests to study the behavior of the proposed debiasing method to estimate the drift of the GBM. Since the default probability is a widely used metric for risk management, we will also analyze how removing the bias on $\widehat{\mu}$ helps to get a more precise estimate of that probability.

3.2.1 Drift

We first compare the (conditional) expected NMLE, CMLE and debiased CMLE with the true value of μ for various pairs of (A_0, μ). The expected debiased CMLE has been computed by simulation with 100,000 paths (see Appendix A.3 for details about the simulation algorithm[8]) whereas values for the NMLE and CMLE are copied from tables 1 and 3 for convenience. All results are shown in Table 4.

In general, the debiasing procedure corrects the upward bias in $\widehat{\mu}^N$ and the downward bias in $\widehat{\mu}$ that we originally obtained in a finite sample. More interestingly, the debiased CMLE works best when the default probability is large which is exactly when the survival bias matters the most. Over large time horizons, the debiasing removes almost all remaining bias to the contrary of the NMLE or CMLE. These results illustrate numerically the first statistical property of the method i.e. it is first-order unbiased.

We now analyze the precision of the debiased CMLE. In Table 5, we show additional descriptive statistics on the debiased CMLE, namely the standard deviation and a few quantiles (10%, 50% and 90%). As is always the case when estimating the drift of a GBM, there is a large uncertainty on the debiased CMLE of the drift. When $T = 1$, the uncertainty is important, especially when the assets are close to the barrier. When $T = 10$, the debiased CMLE is much more precise: the remaining bias almost disappears whereas both the standard deviation and the 10th-90th quantile range decrease significantly.

---

[8]Since the debiased CMLE only depends upon z_0 and z_n, it would not be necessary to simulate the entire path. However, when σ is unknown as will be the case later, the entire path is needed.

		1-year horizon				10-year horizon			
A_0	μ	NMLE	CMLE	DebCMLE	Def prob	NMLE	CMLE	DebCMLE	Def prob
110	-0.1	0.2739	-0.4668	-0.0792	0.8534	0.1089	-0.2883	-0.1064	0.9936
200	-0.1	-0.0642	-0.2300	-0.0662	0.0581	0.0528	-0.2815	-0.1006	0.8989
300	-0.1	-0.0987	-0.1164	-0.0781	0.0013	0.0183	-0.2697	-0.1017	0.7689
110	0	0.3127	-0.3035	0.0253	0.7854	0.1370	-0.1167	0.0063	0.9567
200	0	0.0198	-0.0861	0.0320	0.0292	0.0848	-0.1086	0.0066	0.6271
300	0	0.0005	-0.0085	0.0162	0.0004	0.0565	-0.0950	0.0080	0.4010
110	0.1	0.3574	-0.1458	0.1235	0.7056	0.1812	0.0400	0.1102	0.8571
200	0.1	0.1101	0.0460	0.1318	0.0135	0.1365	0.0482	0.1106	0.2836
300	0.1	0.1001	0.0958	0.1110	0.0001	0.1182	0.0603	0.1093	0.1146
110	0.2	0.4085	0.0055	0.2318	0.6189	0.2454	0.1755	0.2084	0.7165
200	0.2	0.2047	0.1678	0.2266	0.0057	0.2114	0.1816	0.2081	0.0841
300	0.2	0.2000	0.1980	0.2082	0.0000	0.2035	0.1887	0.2073	0.0181
110	0.3	0.4664	0.1498	0.3310	0.5309	0.3265	0.2916	0.3047	0.5825
200	0.3	0.3020	0.2817	0.3197	0.0022	0.3027	0.2950	0.3042	0.0195
300	0.3	0.3000	0.2991	0.3054	0.0000	0.3004	0.2978	0.3035	0.0019

Table 4: Conditional expected value of $\widehat{\mu}^N$ (NMLE), $\widehat{\mu}$ (CMLE) and $\widetilde{\mu}$ (DebCMLE) for a 1-year (left) and 10-year horizon (right). Other parameters: $\sigma = 0.3, L = 100$. "Def prob" is short for default probability.

		1-year horizon					10-year horizon				
A_0	μ	Stdev	10%	50%	90%	90-10	stdev	10%	50%	90%	90-10
110	-0.1	1.2296	-0.6384	0.0853	0.4745	1.1129	0.6294	-0.3379	-0.0128	0.1266	0.4645
200	-0.1	0.7226	-0.4231	-0.0218	0.3092	0.7323	0.509	-0.3247	-0.0159	0.114	0.4388
300	-0.1	0.2866	-0.4213	-0.0859	0.287	0.7083	0.5468	-0.3025	-0.0231	0.098	0.4005
110	0	1.0461	-0.5004	0.1563	0.5366	1.037	0.4222	-0.1752	0.0583	0.1847	0.3598
200	0	0.4435	-0.3122	0.0549	0.3998	0.712	0.3513	-0.162	0.0542	0.1709	0.3329
300	0	0.286	-0.3371	0.0088	0.3845	0.7216	0.2832	-0.1449	0.0469	0.1571	0.302
110	0.1	0.9901	-0.3649	0.2301	0.6052	0.97	0.178	-0.0369	0.1336	0.2521	0.289
200	0.1	0.3243	-0.2089	0.1392	0.4914	0.7002	0.1653	-0.0245	0.1297	0.242	0.2664
300	0.1	0.2895	-0.2534	0.1039	0.4873	0.7407	0.199	-0.0119	0.1232	0.2326	0.2445
110	0.2	0.8043	-0.2248	0.3115	0.6804	0.9052	0.1234	0.0807	0.2164	0.3355	0.2547
200	0.2	0.307	-0.1113	0.2254	0.5875	0.6987	0.1153	0.0898	0.2129	0.3276	0.2379
300	0.2	0.2906	-0.1639	0.2036	0.5845	0.7484	0.0972	0.0951	0.2081	0.3233	0.2282
110	0.3	0.633	-0.1031	0.3899	0.7581	0.8612	0.0984	0.1831	0.3056	0.4274	0.2443
200	0.3	0.2945	-0.0265	0.3147	0.6859	0.7123	0.0938	0.187	0.3041	0.4238	0.2368
300	0.3	0.294	-0.0715	0.302	0.6845	0.756	0.0924	0.1877	0.302	0.422	0.2343

Table 5: Descriptive statistics on the debiased CMLE, conditional upon survivorship. Other parameters: $\sigma = 0.3, L = 100$. "stdev" is short for standard deviation, $x\%$ is the x-th quantile, whereas 90-10 is the difference between the 90-th and the 10-th quantile.

20

		1-year horizon					10-year horizon				
A_0	μ	Stdev	10%	50%	90%	90-10	stdev	10%	50%	90%	90-10
110	-0.1	1.5005	-1.4074	-0.1686	0.3844	1.7918	0.6969	-0.7184	-0.1381	0.0783	0.7966
200	-0.1	0.7793	-0.7742	-0.1164	0.2919	1.0662	0.6314	-0.6905	-0.1353	0.0684	0.7589
300	-0.1	0.3482	-0.517	-0.1009	0.2859	0.8029	0.6128	-0.6427	-0.1342	0.0575	0.7002
110	0	1.288	-1.1598	-0.0598	0.4627	1.6225	0.4722	-0.4225	-0.023	0.1562	0.5787
200	0	0.6388	-0.5754	-0.0122	0.3899	0.9653	0.4364	-0.3929	-0.021	0.1452	0.5381
300	0	0.3245	-0.4007	0.0008	0.384	0.7847	0.3822	-0.3514	-0.0191	0.1368	0.4882
110	0.1	1.1438	-0.9199	0.0498	0.5463	1.4662	0.2703	-0.1788	0.0881	0.2379	0.4167
200	0.1	0.4547	-0.4007	0.0949	0.4859	0.8866	0.2431	-0.1497	0.0904	0.231	0.3808
300	0.1	0.3145	-0.2942	0.0997	0.4871	0.7813	0.2311	-0.115	0.0924	0.2257	0.3407
110	0.2	0.9607	-0.677	0.166	0.6351	1.3121	0.1677	0.0113	0.1957	0.3298	0.3185
200	0.2	0.3997	-0.2469	0.1974	0.5847	0.8316	0.1477	0.0332	0.1971	0.3243	0.2911
300	0.2	0.305	-0.1882	0.2016	0.5845	0.7726	0.1237	0.0533	0.198	0.3218	0.2685
110	0.3	0.8026	-0.4719	0.2733	0.7237	1.1956	0.1137	0.1542	0.2978	0.4254	0.2712
200	0.3	0.359	-0.1229	0.298	0.6845	0.8075	0.1048	0.1656	0.2994	0.423	0.2574
300	0.3	0.3032	-0.0852	0.3011	0.6845	0.7696	0.0997	0.1743	0.2999	0.4218	0.2475

Table 6: Descriptive statistics on the CMLE, conditional upon survivorship. Other parameters: $\sigma = 0.3, L = 100$. "stdev" is short for standard deviation, $x\%$ is the x-th quantile, whereas 90-10 is the difference between the 90-th and the 10-th quantile.

We also provide similar statistics for the standard CMLE. These descriptive statistics have been obtained by simulation on the same 100,000 paths than in Table 5.

When we compare the distribution of the debiased CMLE (Table 5) with the standard CMLE (Table 6), we observe that the debiasing not only successfully removes the survivorship bias, but it also makes the estimator more precise. Indeed, in all the cases investigated, the debiased CMLE has much lower bias and also inferior standard deviation and 10th-90th quantile range.

It is also interesting to compare the standard deviation of $\widetilde{\mu}$ or $\widehat{\mu}$ with the base case where assets are always observed, no matter if they reach the barrier or not. In the latter case, there is no such survivorship bias and the standard error is simply $\sigma/\sqrt{T}$ i.e. 0.3 with $T = 1$ and 0.095 with $T = 10$. Comparing these values with what we observe in Table 5 and 6, we see that survivorship bias adds significant uncertainty on the estimator of the drift, uncertainty that increases with the default probability.

21

A_0	μ	1-yr DP	1-year horizon					10-year horizon				
			Mean	Stdev	10%	50%	90%	Mean	Stdev	10%	50%	90%
110	-0.1	0.8534	0.7037	0.2230	0.3894	0.7179	0.9951	0.8051	0.0992	0.6830	0.7949	0.9559
200	-0.1	0.0581	0.1085	0.1909	0.0020	0.0342	0.3000	0.0911	0.1739	0.0120	0.0328	0.1993
300	-0.1	0.0013	0.0123	0.0494	0.0000	0.0011	0.0256	0.0210	0.1145	0.0001	0.0006	0.0096
110	0	0.7854	0.6587	0.2309	0.3453	0.6573	0.9851	0.7517	0.1005	0.6324	0.7400	0.8949
200	0	0.0292	0.0737	0.1526	0.0008	0.0193	0.1882	0.0493	0.1156	0.0074	0.0194	0.0852
300	0	0.0004	0.0064	0.0301	0.0000	0.0004	0.0130	0.0077	0.0674	0.0001	0.0002	0.0021
110	0.1	0.7056	0.6089	0.2360	0.3013	0.5923	0.9626	0.6863	0.0968	0.5729	0.6769	0.8121
200	0.1	0.0135	0.0465	0.1135	0.0003	0.0097	0.1114	0.0216	0.0605	0.0039	0.0105	0.0349
300	0.1	0.0001	0.0033	0.0185	0.0000	0.0001	0.0062	0.0017	0.0310	0.0000	0.0001	0.0005
110	0.2	0.6189	0.5512	0.2359	0.2584	0.5210	0.9176	0.6093	0.0898	0.5005	0.6044	0.7217
200	0.2	0.0057	0.0276	0.0808	0.0001	0.0045	0.0625	0.0081	0.0218	0.0017	0.0051	0.0146
300	0.2	0.0000	0.0016	0.0111	0.0000	0.0000	0.0026	0.0002	0.0098	0.0000	0.0000	0.0001
110	0.3	0.5309	0.4950	0.2317	0.2200	0.4552	0.8553	0.5286	0.0824	0.4252	0.5261	0.6338
200	0.3	0.0022	0.0162	0.0572	0.0000	0.0019	0.0354	0.0031	0.0045	0.0006	0.0021	0.0064
300	0.3	0.0000	0.0008	0.0071	0.0000	0.0000	0.0010	0.0000	0.0013	0.0000	0.0000	0.0000

Table 7: Descriptive statistics on the estimator of the 1-year default probability computed with the debiased CMLE. Other parameters: $\sigma = 0.3, L = 100$. 1-yr DP is short for the true 1-year default probability, "stdev" is short for standard deviation and $x\%$ is the x-th quantile. Any value below 0.00005 is shown as 0.0000 and any value over 0.99995 is shown as 1.0000.

3.2.2 Default probability

In various applications in risk management, the drift is an important input for various metrics used for decision-making such as the default probability. In this section, we are interested in the computation of the 1-year default probability, using $\widehat{\mu}$ or $\widetilde{\mu}$ estimated with $T = 1$ or $T = 10$. Results are shown in tables 7 and 8 for the debiased CMLE and the CMLE respectively.

Because it depends on the drift which has a lot of embedded uncertainty, the default probability is difficult to accurately estimate. Take for example a company with $A_0 = 200$ and $\mu = 0.1$ whose true default probability is 1.4%. Its S&P rating would be between BB and B. With $T = 10$ to estimate the drift, the debiased CMLE yields a corresponding estimate of the default probability between 0.4% and 3.5% (with a confidence level of 80%) with a mean of 2.2%. Interestingly, the corresponding range of default probability with the CMLE is 0.4% to 8.0% with a mean of 4.3%. The range is more than twice as large and an important bias

A_0	μ	1-yr DP	1-year horizon					10-year horizon				
			Mean	Stdev	10%	50%	90%	Mean	Stdev	10%	50%	90%
110	-0.1	0.8534	0.8048	0.2167	0.4597	0.8916	1.0000	0.8667	0.1021	0.7237	0.8755	0.9976
200	-0.1	0.0581	0.2053	0.2894	0.0024	0.0645	0.7327	0.1930	0.2664	0.0174	0.0726	0.6358
300	-0.1	0.0013	0.0266	0.0968	0.0000	0.0013	0.0510	0.0616	0.1937	0.0002	0.0019	0.1109
110	0	0.7854	0.7604	0.2352	0.3981	0.8277	1.0000	0.8066	0.1144	0.6574	0.8022	0.9743
200	0	0.0292	0.1432	0.2442	0.0009	0.0319	0.4886	0.1058	0.1936	0.0092	0.0340	0.2668
300	0	0.0004	0.0135	0.0624	0.0000	0.0004	0.0218	0.0243	0.1213	0.0001	0.0005	0.0146
110	0.1	0.7056	0.7083	0.2513	0.3388	0.7469	0.9997	0.7277	0.1174	0.5854	0.7156	0.8967
200	0.1	0.0135	0.0925	0.1934	0.0003	0.0141	0.2751	0.0434	0.1104	0.0043	0.0146	0.0792
300	0.1	0.0001	0.0066	0.0396	0.0000	0.0001	0.0089	0.0056	0.0565	0.0000	0.0001	0.0016
110	0.2	0.6189	0.6452	0.2622	0.2835	0.6488	0.9965	0.6337	0.1086	0.5053	0.6227	0.7769
200	0.2	0.0057	0.0552	0.1440	0.0001	0.0058	0.1365	0.0132	0.0420	0.0017	0.0058	0.0228
300	0.2	0.0000	0.0031	0.0241	0.0000	0.0000	0.0033	0.0006	0.0170	0.0000	0.0000	0.0002
110	0.3	0.5309	0.5808	0.2668	0.2363	0.5542	0.9817	0.5394	0.0933	0.4268	0.5329	0.6591
200	0.3	0.0022	0.0320	0.1044	0.0000	0.0023	0.0672	0.0039	0.0103	0.0006	0.0022	0.0077
300	0.3	0.0000	0.0014	0.0143	0.0000	0.0000	0.0011	0.0000	0.0036	0.0000	0.0000	0.0000

Table 8: Descriptive statistics on the estimator of the 1-year default probability computed with the CMLE. Other parameters: $\sigma = 0.3, L = 100$. 1-yr DP is short for the true 1-year default probability, "stdev" is short for standard deviation and $x\%$ is the x-th quantile. Any value below 0.00005 is shown as 0.0000 and any value over 0.99995 is shown as 1.0000.

remains.

We know that the CMLE always underestimates the drift and should therefore overestimate the default probability. Whenever the true 1-year default probability is much lower than 1 (say e.g. below 60%), this is exactly what we observe in tables 7 and 8, that is the debiased CMLE provides a better performance than the standard CMLE with a lower bias on the 1-year default probability whereas the debiasing yields estimates that are also more precise (smaller standard deviation and tighter quantiles).

When the true 1-year default probability is very high (say above 60%), the CMLE generates highly negative drifts that often result in 100% default probabilities, no matter how negative the CMLE is. Simulated default probabilities that are bounded to ≈ 1 contribute to reducing the bias and uncertainty for the estimator of the 1-year default probability calculated from the CMLE. As a result, in a few instances where the default probability is very high, the CMLE has a lower bias than the debiased CMLE but this slight advantage quickly vanishes when $T = 10$.

23

Finally, it would be possible to further improve the results obtained in Table 7 by applying the debiasing on the 1-year default probability, which is a function $\widehat{p(\mu)}$, rather than on $\widehat{\mu}$.

4 Credit risk with survivorship

There are generally two classes of default triggering mechanisms in structural credit risk models. The first one, present in Merton-like models, assumes that default only occurs at maturity if A_T is insufficient to repay the debt. The second mechanism, associated to first-passage models, supposes that default occurs when A_t attains a solvency barrier L. By construction, only the second type of model is subject to the survivorship bias.

A structural credit risk model specifies the capital structure of the firm, linking the firm's asset value to its equity. In such model, equity represents an option on the firm's assets. Whereas Merton-like models rely on a European call option to link equity values to the firm's asset prices, first-passage models do so with down-and-out call options.

One difficulty that arises with structural models is the inability of an investor to observe the market value of a firm's assets. Therefore, inference of structural models is often carried out with stock prices rather than directly from the firm's assets.

In this section, we relax two important assumptions: we suppose $\underline{A}$ is unobserved and has to be inferred from stock prices; we assume σ is unknown and estimate it as well. The first part of this section presents the credit risk model that we use in this paper. The second part explains parameter estimation from stock prices using an extended version of Vassalou and Xing (2004) that accounts for survivorship.

4.1 Model specification

Brockman & Turtle (2003) value the outstanding equity of a firm in a framework where $\underline{A}$ is a GBM and default is triggered as soon as A_t attains a solvency barrier $L > A_0$ (as defined in Black & Cox (1976)). They assume that at each point in time, the firm's capital structure is composed of equity S and liabilities D. Liabilities consist of not only corporate bonds but also of any other form of payable account. If the firm defaults, all of the asset value is transferred to debt holders.

Second, it is assumed that the barrier L is equal to the nominal value of liabilities, thus requiring surviving firms to be able to fulfill all their obligations during the observation period. It is necessary to specify the barrier L exogenously since it cannot be estimated from asset values nor equity prices.

With a constant risk-free rate r and using standard hedging arguments, the value of a unit of equity corresponds to the value of a down-and-out call option on the firm's assets. This option has a strike price equal to the total liability value D, a knock-out barrier of L, and a maturity equal to T. Thus, within this framework, the value of equity is:

$$
\begin{aligned}
S_t \quad &:= \quad \mathcal{S}\left(A_t; \theta\right) \\
&= \quad A_t \Phi(a_t) - D e^{-r(T-t)} \Phi\left(a_t - \sigma\sqrt{T-t}\right) \\
&\quad - A_t \left(\frac{L}{A_t}\right)^{2\eta} \Phi\left(b_t\right) \\
&\quad + D e^{-r(T-t)} \left(\frac{L}{A_t}\right)^{2\eta-2} \Phi\left(b_t - \sigma\sqrt{T-t}\right),
\end{aligned}
\tag{4.1}
$$

where

$$
\begin{aligned}
a_t \quad &= \quad \frac{\ln A_t - \ln D + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \\
b_t \quad &= \quad \frac{2\ln L - \ln A_t - \ln D + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \\
\eta \quad &= \quad \frac{1}{\sigma^2}\left(r + \frac{\sigma^2}{2}\right).
\end{aligned}
$$

Note that the first two terms of Equation (4.1) correspond to the equity value in the Merton (1974) model. The third and fourth terms, which depend on the debt ratio L/A_t, can be viewed as terms that correct Merton's equity value for early default.

25

4.2 Parameter estimation

Assume now that $\underline{A}$ is not directly observed but as an investor, we only observe $\underline{S} = \{S_t, t \geq 0\}$ at discrete time points $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$ with $h = t_i - t_{i-1} = T/n, \forall i = 1, 2, \ldots, n$. Provided that we know all model parameters and there are no market imperfections[9] on stock prices, then $S_t = \mathcal{S}(A_t)$ is a monotone increasing function of A_t. To recover A_t, it only suffices to invert $\mathcal{S}$. As expected, $\underline{S}$ is affected by the survivorship bias as well due to $\underline{A}$.

Suppose that D, L, T and r are known so that the goal is to use observed equity prices $S_0, S_{t_1}, S_{t_2}, \ldots, S_{t_n}$ to estimate μ and σ. We follow Vassalou and Xing (2004) and Bharath and Shumway (2008) and adopt a two-step approach to estimate these parameters.

First of all, we know from Section 2.2 that σ is practically unaffected by the survivorship bias. Moreover, we know from Duan et al. (2005) that the MLE of Duan (1994) and the iterative procedure of Vassalou and Xing (2004) are equivalent. Therefore, even in the presence of survivorship, we can determine σ using the exact same iterative procedure described by Vassalou and Xing (2004).

For the purpose of the method, let $\widehat{\sigma}^{(k)}$ be the estimate of σ at step k and let $\underline{A}^{(k)}$ be the asset value process such that $A_t = \mathcal{S}^{-1}(S_t; \widehat{\sigma}^{(k)})$ i.e. the asset value obtained by inverting the equity price function using $\widehat{\sigma}^{(k)}$. The algorithm goes as follows:

1. Using an initial estimate $\widehat{\sigma}^{(0)}$ and observed equity prices $\underline{S}$, we back out a time series of asset values $\underline{A}^{(0)}$.

2. Then, $\widehat{\sigma}^{(1)}$ is computed as the MLE of σ for a GBM having observations $\underline{A}^{(0)}$ (see Equation 2.3).

3. Using the estimate $\widehat{\sigma}^{(1)}$ and the same observed equity prices $\underline{S}$, we back out another time series of asset values $\underline{A}^{(1)}$.

4. Steps 2 and 3 are repeated until convergence.

The procedure stops when $|\widehat{\sigma}^{(k)} - \widehat{\sigma}^{(k-1)}| < \epsilon$ for ϵ small. Moreover, no estimate of μ is necessary at this point because the equity price is independent from that parameter (equity

---

[9]Also known as trading/microstructure noise, see e.g. Duan & Fulop (2009).

pricing is done under the risk-neutral measure).

Once an estimate of σ is obtained at the k-th iteration, we work with the time series of $\underline{A}^{(k)}$ to estimate μ. Given that μ is affected by the survivorship bias, we apply the debiased MLE introduced in Section 3[10].

A simulation study, available in the Internet Appendix, analyzes the behavior of the debiased CMLE using this two-step procedure. Overall, we find that in this case the debiased CMLE performs similarly to the case where σ was known and $\underline{A}$ was observed (see Section 3.2).

5 Empirical study

The analyses presented in Section 2 have shown the existence of a bias in the estimation of the parameter μ. In this section, we assess the economic impact of ignoring or miscorrecting for the survivorship bias when conducting inference of a first-passage structural credit risk model. To this end, we will use the NMLE, the CMLE and the debiased CMLE to assess the credit risk of thousands of companies. We first present the data employed throughout our study and then we investigate several aspects of the empirical differences tied to each of the three estimates of μ.

5.1 Data, assumptions, and methodology

Our dataset is composed of all firms in the intersection of CRSP[11] and Compustat between January 1979 and December 2016. Sample filters and model inputs are defined following Bharath and Shumway (2008). First, financial companies identified with CUSIP codes 6021, 6022, 6029, 6035, and 6036 are excluded from the sample. Second, model inputs such as face value of debt (D) and barrier (L) are obtained from Compustat annual data. Debt's face value D is defined as "Debt in one year" plus half of "Long-term debt" fields in Compustat. We also use this value for the barrier L because it captures the ability of a firm to service long- and short-term debt, thus indicating the firm's survival capability in the short term.

---

[10]In the empirical analysis we also employ the NMLE and CMLE to assess the bias of these estimates.

[11]Center for Research in Security Prices

The debt maturity T is set to 1 year[12].

The daily equity value is calculated as the price per share from CRSP multiplied by the total number of shares outstanding. Regarding the risk-free rate r, we employ a monthly time series of 1-year Treasury maturity rates provided by the Board of Governors of the Federal Reserve system[13].

Our study is conducted as follows. For each firm and month in our sample, we extract the previous 12 months of daily equity values along with the current estimate of the debt's face value, barrier, and risk-free rate to compute estimates of σ and μ as discussed in Section 4.2. In the iterative procedure of Section 4.2, we suppose the algorithm has attained convergence when $\epsilon = 10^{-4}$ with $\widehat{\sigma}^{(0)}$ computed as the sample standard deviation of daily log equity returns[14]. Moreover, we discard a month if a) there are less than 200 daily equity observations or b) the value of assets was below the debt at any point in the twelve month period. Our final sample consists of 13,794 firms and a total number of 1,231,167 firm-months observations with complete data.

5.2 Distribution of μ

We seek to compare the NMLE, CMLE, and debiased CMLE of μ across companies. For each given company in the sample, we first compute the median estimate of μ over time. Then, we use these estimates to compute percentiles across all firms in the sample for each of the methods under consideration. By comparing different percentiles across companies, we are able to better characterize the bias across companies with different solvency profiles. To conduct statistical analyses, we also report confidence intervals for percentiles.

Table 9 reports percentiles across companies for the estimates of μ based upon the NMLE,

---

[12]The choice of maturity in this manuscript was done to compare our results with papers that analyze default probabilities over a one year period. Papers such as Vassalou and Xing (2004) and Bharath and Shumway (2008) use a debt maturity of one year, thus looking at the ability of a firm to survive with the current debt level over a short period of time. Nonetheless, the literature is not conclusive about which debt maturity should be used empirically. For instance, in the context of barrier-based structural models, Dionne and Laajimi (2012) use a 20-year debt, while Brockman and Turtle (2003) and Reisz and Perlich (2007) use 10 years.

[13]Available at http://research.stlouisfed.org/fred/data/irates/gs1

[14]If one is interested in conducting parameter inference, one way to compute standard errors is to use e.g. bootstrap methods. Given that our objective is to empirically assess the point estimates of the three methods, we do not compute these errors.

	DebCMLE			CMLE			NMLE		
Percentile	LB	μ (%)	UB	LB	μ (%)	UB	LB	μ (%)	UB
P10	-6.84	-6.19	-5.35	-19.64	-18.38	-17.17	0.33	0.88	1.30
P25	5.80	6.12	6.53	1.11	1.60	2.11	9.66	9.94	10.23
P50	16.45	16.82	17.12	12.98	13.37	13.76	19.65	19.99	20.35
P75	34.35	35.32	36.46	28.68	29.48	30.18	36.15	37.00	37.77
P90	66.99	69.24	71.44	56.34	57.73	59.47	64.59	66.73	68.84

Table 9: Estimates of μ (in %) using different estimation procedures

For each company in the sample, the median estimate of μ is computed over time. This table reports percentiles of the latter across companies for the NMLE, CMLE and debiased CMLE methods. Columns DebCMLE, CMLE and NMLE represents estimates of μ based upon the debiased conditional MLE, conditional MLE and naive MLE respectively. The lower and upper bounds for the 95% confidence intervals are reported as LB and UB. The sample spans 1979 through 2016 and contains 13,794 firms.

CMLE and debiased CMLE (DebCMLE). For both the CMLE and NMLE estimates, there is compelling empirical evidence of an important difference between estimates: percentiles are systematically below or above from those computed with DebCMLE. The strongest results are observed for the CMLE, in which all percentiles are statistically below debiased estimates.

Although clearly above the DebCMLE for most of the percentiles, the NMLE is not significantly different from the DebCMLE for large values of μ. This result hinges on the fact that for large values of μ, the default probability is small and so is the resulting survivorship bias; the correction needed in this case is insignificant. Nonetheless, the CMLE in these cases still exhibits a significant downward bias. Overall, these results suggest that the biases documented in Section 2 are present empirically and can lead to significant differences in the estimates of the drift and, in turn, of the default probability.

5.3 Risk profiles

The previous analysis has shown that for most companies, there can be important differences in estimates of μ across the three methods. We now want to study how these differences impact the risk profile of each company. To this end, we group firms according to their default probability (which depends directly on estimates of μ) and compare several characteristics (such as asset volatility, leverage, etc.) across these groups. If all estimation methods produce the same risk profiles, then a group's characteristics will be the same for each of the three methods.

The creation of groups and computation of their characteristics is done as follows. For each estimation method, we compute the 1-year default probability for each month in the sample using Equation 2.2. Next, we compute the median of this monthly series across time. In a similar fashion, firm characteristics are computed using median values of the following monthly series:

- Asset return μ;

- Asset return volatility σ;

- Leverage ratio (defined as the ratio between the face value of debt (D) and the asset value (A_t))[15];

- Risk-adjusted measure of performance defined as the ratio of μ over σ;

Companies are then sorted by deciles according to their median default probability (over months). Then, a given characteristic for a decile is calculated as the median of the characteristic across companies within that decile.

Table 10 shows several company characteristics by decile group. In the table, firms in decile 1 have the lowest default probability estimates whereas firms in decile 10 have the largest. Observe that as default probabilities increase (from decile 1 to 10), the leverage and asset volatility increase as well. Given that we observe these monotonic relationships across all three methods, we can conclude that the level and not the ordering of risk profiles is impacted when estimates of μ do not properly account for the survivorship bias. When we look at risk-adjusted asset returns (columns Performance), it becomes clear that controlling for the survivorship bias is important in the analysis of companies with high degree of default risk (deciles 6 to 10). For example in the last decile, the risk-adjusted return is still positive when survivorship is ignored whereas it is largely negative when the survivorship bias is miscorrected with the CMLE. Using the debiased CMLE, it is true that these firms underperform but clearly not at the extent suggested by the CMLE.

---

[15]We use the last asset value of $\underline{A}^{(k)}$ in the iterative procedure of Section 4.2.

	μ (%)			σ(%)			$Leverage$(%)			$Performance$		
Decile	DebCMLE	CMLE	NMLE	DebCMLE	CMLE	NMLE	DebCMLE	CMLE	NMLE	DebCMLE	CMLE	NMLE
1	19.08	18.88	19.59	34.19	34.19	34.19	1.79	1.79	1.79	0.62	0.61	0.62
2	18.58	17.90	19.37	35.80	35.80	35.80	5.84	5.84	5.84	0.59	0.57	0.61
3	19.66	18.82	21.21	37.64	37.64	37.64	8.52	8.52	8.52	0.58	0.55	0.60
4	19.73	18.36	21.40	40.63	40.63	40.63	10.06	10.06	10.06	0.54	0.50	0.59
5	17.73	16.31	21.13	41.50	41.92	38.47	13.11	13.04	14.96	0.48	0.44	0.60
6	17.14	14.81	22.34	45.25	45.20	45.67	15.39	15.30	15.09	0.46	0.40	0.57
7	18.05	13.78	24.02	48.46	48.67	48.17	18.35	18.13	18.48	0.45	0.35	0.59
8	15.19	8.04	21.27	54.42	54.55	54.12	21.71	21.24	21.76	0.36	0.22	0.49
9	11.44	-0.01	16.82	57.76	58.99	59.15	28.07	27.14	27.02	0.25	0.00	0.35
10	-3.47	-33.25	7.41	66.97	63.11	69.79	39.63	41.66	37.61	-0.09	-0.67	0.13

Table 10: Parameter characteristics by default probability deciles

Firms are ordered by default probabilities and grouped by deciles. For each given company and estimation method, the median 1-year default probability is computed over time and the same goes for various firm characteristics such as asset return (μ), asset return volatility (σ), leverage and the risk-adjusted performance of the firm's assets. Values reported in this table are median risk characteristic computed within each group. Our sample spans 1979 through 2016 and contains 13,794 firms.

5.4 Effect on default probabilities

We now explore the survivorship bias effect on the default probability of a firm over time. Specifically, we focus on risky companies across recession periods (NBER business cycle contraction dates). Each month, we compute the 90th percentile of the default probability across firms. We look at the time series of this quantity to determine how the bias shifts the tails of the distribution in periods of high economic uncertainty.

Each panel of Figure 2 displays these percentiles for four recessions: 1981/07-1982/11, 1990/07-1991/03, 2001/03-2001/11 (burst of the dot-com bubble) and 2007/12-2009/06 (Great Recession). Due to the overestimation (underestimation) of μ with the NMLE (CMLE), we observe that the default probability is systematically lower (higher) than the debiased estimates.

The overall picture that emerges from these time series is that the solvency of the riskiest companies (as measured by the 90th percentile) is largely affected by ignoring or miscorrecting for the survivorship bias. At the worst of the 1990-1991 recession, the 90th quantile of the default probability computed with the DebCMLE is 60%, which is largely overstated by the CMLE (85%) and significantly understated by the NMLE (45%). Similar differences are

31

(a) July 1981 - November 1982

(b) July 1990 - March 1991

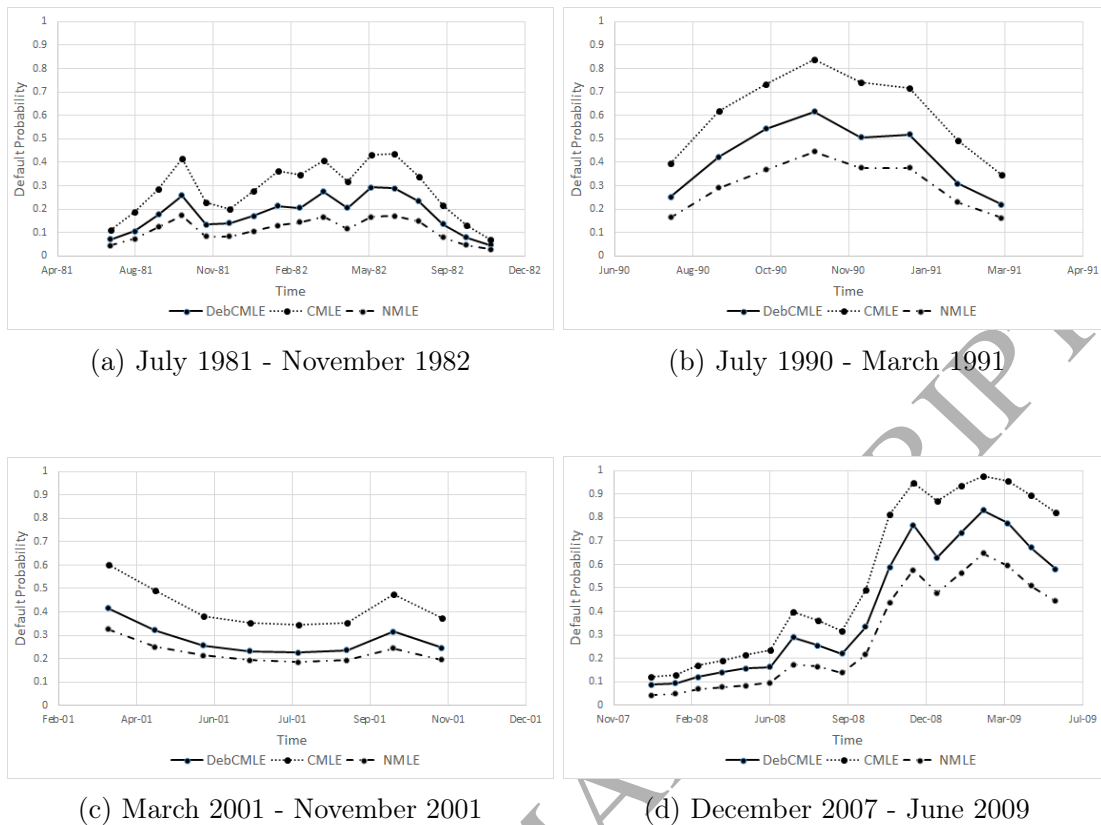(c) March 2001 - November 2001

(d) December 2007 - June 2009

Figure 2: Evolution of the default probability of the 90-th percentile company across NBER recession periods

This figure shows the monthly 90-th percentile of the default probability across firms using three different estimates of μ. DebCMLE represents the estimate of μ computed with the debiased conditional MLE. CMLE stands for the conditional MLE of μ whereas NMLE is the estimate of μ that ignores survivorship. Recession periods, identified below each graph, are defined by NBER as periods of business cycle contractions.

observed at the peak of the recent financial crisis — around March 2009. In brief, the gap in these default probabilities varies over time and increases in periods of recession.

5.5 Misclassification of companies

From a risk management perspective, it is important to quantify the proportion of companies over- or under-classified as risky ones. Every year and for each estimation method, we compute the proportion of firms with an average 1-year default probability above a predefined threshold value. Using the debiased CMLE of μ, this threshold is set to the 95-th and 99-th percentile of the default probability across all companies and months in the sample. Figure 3 displays the annual proportion of companies having a default probability higher than this
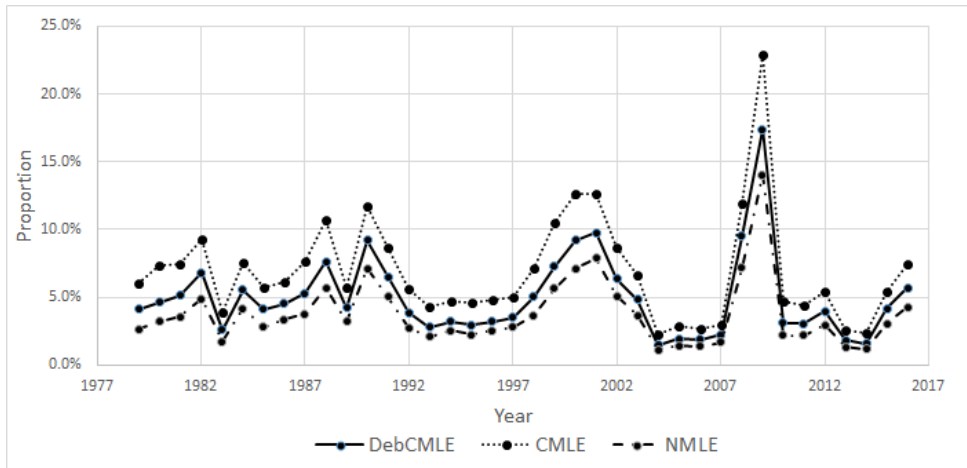
threshold.

Panels (a) and (b) show that these proportions have large variability over time with important spikes during economic recessions, in line with our previous results. More importantly, the proportion of misclassified risky firms can be systematically different depending on the estimation method: CMLE provides a larger proportion of risky companies whereas the NMLE a lower number. Given the previous evidence, these differences should not be surprising; however, they show the economic value of accounting for the survivorship bias.

For instance, in Panel A (95% threshold) during the year 2000 the NMLE classifies 7.1% of companies as risky, while the CMLE does so for 12.7% of the sample. In contrast, the proportion that comes from the debiased procedure is 9.2%. These differences become more important when looking at very risky companies (Panel B of in Figure 3 i.e. 99% threshold), in which case the proportions can differ more substantially. At the worst of the Great Recession, the estimate of the true proportion of firms above the 99% threshold is about 5% with the debiasing procedure. In contrast, only 1% of companies are classified as very risky when survivorship is ignored, and 9% receive this classification with the CMLE.

6 Conclusion

In this paper, we have studied the maximum likelihood estimation of the GBM conditional on survival. Our results show the existence of an important bias in the drift when (1) we ignore survival of the firm (naive MLE) or (2) when we condition on survival without correcting for the large negative finite sample bias (conditional MLE). Therefore, we proposed a debiasing procedure that is first-order unbiased, second-order efficient, and holds for non-iid samples.

We found that the theoretical biases in the drift translate empirically and economically into important biases in default probabilities. Identifying the proportion of risky companies is an important task for many financial institutions since capital requirements depend on these classifications. Ignoring or miscorrecting for the survivorship bias thus has significant economic consequences for institutional investors and stresses the importance of correctly estimating μ for risk management purposes. The fact that these biases expand during periods of financial distress is relevant for internal risk assessment of solvency models. It still re-

33

(a) Threshold level at 95%
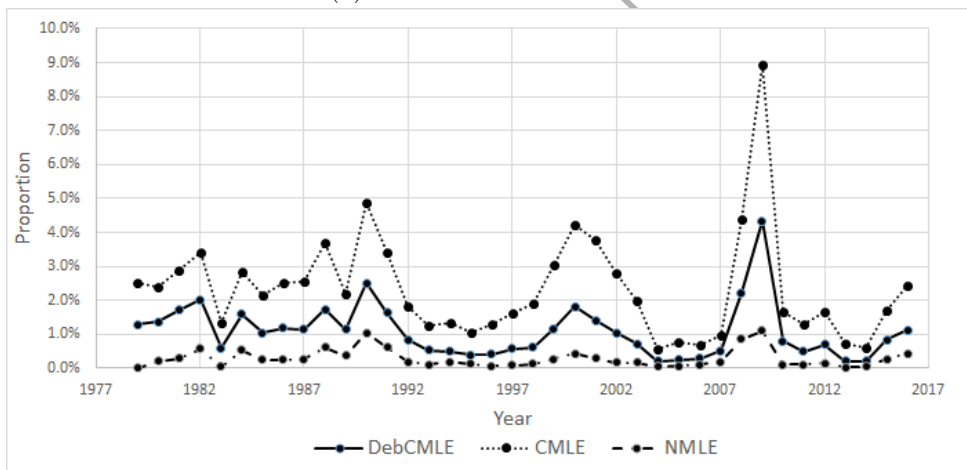
(b) Threshold level at 99%

Figure 3: Proportion of riskiest firms according to the estimate employed for μ.
This figure shows the proportion of companies with annual default probabilities higher than a given threshold. The threshold is set to the 95-th or 99-th percentile of the default probability across all companies and months computed with the debiased CMLE. With default probabilities computed across firms and methods, each panel shows the proportion of firms having a default probability higher than the threshold for each of the three estimation methods. DebCMLE represents the estimate of μ computed with the debiased conditional MLE. CMLE stands for the conditional MLE of μ whereas NMLE is the estimate of μ that ignores survivorship. The sample spans 1979 through 2016 and contains 13,794 firms.

34

mains to determine if appropriately correcting for the survivorship bias materially improves prediction of bankruptcies.

There are additional areas in which this research can be expanded in the future. For example, we have used a structural credit risk model that requires the default barrier to be known or exogenously estimated. It would then be interesting to analyze the effects of survivorship when the default barrier is a latent random variable or stochastic process. Furthermore, the recent literature in the area of credit risk model estimation increasingly relaxes the assumption that stock prices and other credit-sensitive security prices are perfectly observed. For example, Duan & Fulop (2009) have shown that ignoring trading noise overestimates the asset volatility. Such trading/microstructure noise should affect the survivorship bias in a non-trivial manner.

References

[1] Afik, Z., O. Arad, & K. Galil (2016), "Using Merton model for default prediction: An empirical assessment of selected alternatives", *Journal of Empirical Finance* 35, 43-67.

[2] Bauer, M. D., G. D. Rudebusch, & J. C. Wu (2012), "Correcting estimation bias in dynamic term structure models", *Journal of Business and Economic Statistics* **30**, 454-467.

[3] Black, F. and J.C. Cox (1976), "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions", *Journal of Finance* **31**, 351-367.

[4] Bharath, S. T., & T. Shumway (2008), "Forecasting default with the Merton distance to default model", *Review of Financial Studies* **21**, 1339-1369.

[5] Brockman, P. and H. J. Turtle (2003), "A barrier option framework for corporate security valuation", *Journal of Financial Economics* **67**, 511-529.

[6] Brown, S. J., W. Goetzmann, R. G. Ibbotson & S. A. Ross (1992), "Survivorship bias in performance studies", *Review of Financial Studies* **5**, 553-580.

[7] Brown, S. J., W. N. Goetzmann, & S. A. Ross (1995), "Survival", *Journal of Finance* **50**, 853-873.

[8] Cox, D. R., & D. V. Hinkley (1974), "Theoretical statistics", CRC Press.

[9] Dionne, G., & S. Laajimi (2012), "On the determinants of the implied default barrier", *Journal of Empirical Finance* **19**, 395-408.

[10] Duan, J.-C. (1994), "Maximum Likelihood Estimation using Price Data of the Derivative Contract", *Mathematical Finance* **4**, 155-167.

[11] Duan, J.-C. (2000), "Maximum Likelihood Estimation using Price Data of the Derivative Contract: Correction", *Mathematical Finance* **10**, 461-462.

[12] Duan, J.-C. & A. Fulop (2009), "Estimating the Structural Credit Risk Model When Equity Prices Are Contaminated by Trading Noises", *Journal of Econometrics* **150**, 288-296.

[13] Duan, J.-C., J.-G. Simonato, G. Gauthier, S. Zaanoun (2004), "Estimating Merton's Model by Maximum Likelihood with Survivorship Consideration", EFA 2004 Maastricht Meetings Paper No. 4190, SSRN abstract i.d. 557088.

[14] Duan, J.-C., J.-G. Simonato and G. Gauthier (2005) "On the Equivalence of the KMV and Maximum Likelihood Methods for Structural Credit Risk Models", GERAD papers, G-2005-06.

[15] Elton, E.J., M.J. Gruber and C.R. Blake (1996), "Survivorship Bias and Mutual Fund Performance", *Review of Financial Studies* **9**, 1097-1120.

[16] Ericsson, J. and J. Reneby (2005), "Estimating Structural Bond Pricing Models", *Journal of Business* **78**, 707-735.

[17] Forte, S., & L. Lovreta (2012), "Endogenizing exogenous default barrier models: The MM algorithm", *Journal of Banking & Finance* **36**, 1639-1652.

[18] Jones, E.P., S.P. Mason and E. Rosenfeld (1984), "Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation", *Journal of Finance* **39**, 611-625.

[19] Lehmann, E.L. (1983), "Theory of Point Estimation", Chapman and Hall, New York.

[20] Li, M., N. D. Pearson & A. M. Poteshman (2004), "Conditional estimation of diffusion processes", *Journal of Financial Economics* **74**, 31-66.

[21] Merton, R.C. (1974), "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance* **29**, 449-470.

[22] Merton, R.C. (1980), "On Estimating the Expected Return on the Market: an Explanatory Investigation", *Journal of Financial Economics* **8**, 323-361.

[23] Reisz, A. and C. Perlich (2007), "A Market-Based Framework for Bankruptcy Prediction", *Journal of Financial Stability* **3**, 85-131.

[24] Standard & Poors (2016), "Default, Transition, and Recovery: 2015 Annual Global Corporate Default Study And Rating Transitions", S&P Ratings Direct.

[25] Tang, C. Y., & S. X. Chen (2009), "Parameter estimation and bias correction for diffusion processes", *Journal of Econometrics* **149**, 65-81.

[26] Vassalou, M. and Y. Xing (2004), "Default Risk in Equity Returns", *Journal of Finance* **59**, 831-868.

[27] Wong, H. Y. & T. W. Choi (2009), "Estimating default barriers from market information", *Quantitative Finance* **9**, 187-196.

[28] Yu, J. (2012), "Bias in the estimation of the mean reversion parameter in continuous time models", *Journal of Econometrics* **169**, 114-122.