(Non) Consistency of the Beta Kernel Estimator for Recovery Rate Distribution

C. GOURIEROUX, A. MONFORT

CREST and University of Toronto, CNAM and CREST

28 mars 2007
INTRODUCTION

A feature of recovery rate distributions
The asymmetric beta kernel approach

THE NONCONSISTENCY OF THE BETA KERNEL APPROACH

The limit of the beta-kernel estimator
The normalization problem

IMPROVED BETA KERNEL METHODS

COMPARISON OF THE MICRO-BETA METHOD WITH ALTERNATIVE APPROACHES

CONCLUSION

Summary
A key building block of credit risk modelling is the recovery rate, equal to one minus the loss-given-default (LGD).

By definition of LGD, the recovery rate of the regulator lies between 0 and 1.
In practice the recovery rate distribution shows strictly positive weights close to 100% (full recovery) and frequently close to 0% (total loss).

This creates multimodal distributions, which cannot be well-represented by the standard parametric beta distribution. Renault, Scaillet (2004) advocate the use of nonparametric approach based on asymmetric kernel to account this feature and the restricted domain.
INTRODUCTION
THE NONCONSISTENCY OF THE BETA KERNEL APPROACH
IMPROVED BETA KERNEL METHODS
COMPARISON OF THE MICRO-BETA METHOD WITH ALTERNATIVE APPROACHES
CONCLUSION

A feature of recovery rate distributions
The asymmetric beta kernel approach

C. GOURIEROUX, A. MONFORT

(Non) CONSISTENCY OF THE BETA KERNEL ESTIMATOR FOR RECOVERY RATE DISTRIBUTIONS
A feature of recovery rate distributions

The asymmetric beta kernel approach

The point mass at 100% : the consequence of a truncation of underlying recovery rates larger than 1.

Different reasons to explain why $RR > 1$

i) In the definition of LGD, the 90 days period can be too small, especially for cities or regions

ii) Corporate default can arise when a corporate has transitory difficulties to reimburse, but is structurally in good health.
iii) When computed by the workout process, the RR accounts for penalties, costs associated with collecting

\[ RR = \frac{\text{discounted total payment} - \text{discounted cost}}{EAD} \]

(can be larger than 1, or negative)

iv) For large corporate, the debt is often traded on a secondary market, even after a failure at default time: the market value of the debt is generally smaller than the face value, but it can be reevaluated one month later.
Let us consider a continuous distribution on $[0,1]$. A nonparametric estimator of the density based on a Gaussian kernel can lead to inconsistent results at the boundaries. This asymptotic bias can be eliminated by considering an asymmetric kernel.
The asymmetric beta kernel approach

A beta kernel estimator of the density (Chen (1999))

\[ \hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{n} K(X_i, x/b + 1, (1-x)/b + 1), \]

where

\[ K(u; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} u^{\alpha-1} (1 - u)^{\beta-1}, u \in [0, 1] \]

If \( X_1, \ldots, X_n \) are iid variables with pdf \( f_0 \), if \( n \to \infty, b \to 0 \), then

\[ \hat{f}_n(x) \to f_0(x), \forall x \]

[Bouezmarni, Scaillet (2005)]

but THIS CONVERGENCE IS NOT UNIFORM ON \([0,1]\).
The nonconsistency of the beta kernel approach

The main principles of nonparametric estimation of a distribution function.

i) provide an estimator of the cdf in finite sample, that is, an increasing function satisfying the unit mass property;

ii) provide a consistent estimator of the cdf;

iii) provide a consistent estimator of the density function

For credit risk, focus on the quantile function (VaR), which is the inverse of the cdf.
For the application to recovery rate

point mass at 0 : $p_0$
point mass at 1 : $p_1$

Among the $n$ observations :

$n_0$ are equal to 0
$n_1$ are equal to 1

Continuous part : $f_0(.)$, with probability $1 - p_0 - p_1$
The limit of the beta-kernel estimator

Since $K(0; \alpha, \beta) = K(1; \alpha, \beta) = 0$

\[
\hat{f}_n(x) = \frac{1}{n} \sum_{i : X_i \neq 0, 1} K(X_i, x/b + 1, (1 - x)/b + 1)
\]

\[
= \frac{n - n_0 - n_1}{n} \frac{1}{n - n_0 - n_1} \sum_{i : X_i \neq 0, 1} K(X_i, x/b + 1, (1 - x)/b + 1)
\]

\[
\simeq (1 - p_0 - p_1) f_0(x)
\]

Therefore:

\[
\hat{F}_n(x) \simeq (1 - p_0 - p_1) F_0(x)
\]

significantly differs from the true cdf

\[
F(x) = \begin{cases} 
    p_0 + (1 - p_0 - p_1)F_0(x), & \text{if } 0 \leq x < 1, \\
    1, & \text{if } x = 1.
\end{cases}
\]
The limit of the beta-kernel estimator

- The beta-kernel estimator does not ensure unit mass
- Important effects on the estimated CreditVar

If $p_0 = 6\%$ the true quantile at 5\% is zero; the one deduced from the beta kernel provides a strictly positive RR and an underestimation of the required capital

- A consequence of the non uniform convergence of $\hat{f}$ (the integrals do not necessarily converge)
Examples:

- empirical cdf (historical simulation)
  satisfies 1), 2), not 3)

- Gaussian kernel estimator
  satisfies 1), 2), and 3), except at boundaries

- Beta kernel estimator
  satisfies 3), but not 1) and 2
The general term in the expression of $\hat{f}_n$ is:

$$g_b(x, u) = \frac{\Gamma(2 + 1/b)}{\Gamma(x/b + 1)\Gamma((1 - x)/b + 1)} u^{x/b} (1 - u)^{(1-x)/b}$$

For any fixed $x$, $u \to g_b(x, u)$ defines a pdf.

For any fixed $u$, $x \to g_b(x, u)$ is not a pdf.

Very different from a Gaussian kernel $\varphi(x - u)$, where both partial functions are pdf.
The nonconsistency of the beta kernel approach

Improved beta kernel methods

Comparison of the micro-beta method with alternative approaches

Conclusion

The limit of the beta-kernel estimator

The normalization problem

Figure 1: Kernel as a function of x for u given (b = .2)

u = .1, u = .2, u = .3, u = .4, u = .5
FIGURE 2: Kernel as a function of $u$ for $x$ given $b = 0.2$

$x = 0.1, 0.2, 0.3, 0.4, 0.5$
Focus on the total mass, i.e. the integral with respect to $x$
Improved Beta Kernel Methods

Let us focus on the properties of the kernel estimator, when the distribution is continuous (or for the continuous part of the RR distribution)

Two normalizations can be introduced

- a global normalization (on a portfolio)
- a normalization by corporate.
The basic estimator
\[ \hat{f}_n(x) = \frac{1}{n} \sum_{i=1}^{1} K(X_i, x/b + 1, (1 - x)/b + 1) \]

The estimator with global normalization
\[ \hat{f}_n^1(x) = \frac{\hat{f}_n(x)}{\int_0^1 \hat{f}_n(x) dx}. \]

The estimator with normalization by corporate
\[ \hat{f}_n^2(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{K(X_i, x/b + 1, (1 - x)/b + 1)}{\int_0^1 K(X_i, x/b + 1, (1 - x)/b + 1) dx}. \]
The finite sample properties of the three estimators have been compared by Monte-Carlo.

Three possible true distributions in the beta family

- U-Shape for $\alpha = 0.5, \beta = 0.5$;
- Skewed shape for $\alpha = 2.5, \beta = 0.5$;
- Bell-Shape for $\alpha = 2.5, \beta = 2.5$;

$n = 100$

number of replications : 1000
FIGURE 5: Beta density functions
alpha=.5,beta=.5(solid),alpha=2.5,beta=.5(dashes)
alpha=2.5,beta=2.5(short dashes)
Different mean squared errors are provided

Table 1: Comparison of the three beta kernel methods.

<table>
<thead>
<tr>
<th>Mean Squared Errors; $n = 100$; 1000 replications.</th>
<th>Global</th>
<th>Left Tail</th>
<th>Right Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha = 0.5, \beta = 0.5$ ($U$-shape)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plain Beta</td>
<td>0.279</td>
<td>1.334</td>
<td>1.324</td>
</tr>
<tr>
<td>Macro-Beta</td>
<td>0.254</td>
<td>1.158</td>
<td>1.149</td>
</tr>
<tr>
<td>Micro-Beta</td>
<td>0.135</td>
<td>0.611</td>
<td>0.615</td>
</tr>
<tr>
<td>$\alpha = 2.5, \beta = 0.5$ (Skewed shape)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plain Beta</td>
<td>0.687</td>
<td>0.0013</td>
<td>6.71</td>
</tr>
<tr>
<td>Macro-Beta</td>
<td>0.561</td>
<td>0.0018</td>
<td>5.33</td>
</tr>
<tr>
<td>Micro-Beta</td>
<td>0.217</td>
<td>0.0013</td>
<td>2.02</td>
</tr>
<tr>
<td>$\alpha = 2.5, \beta = 2.5$ (Bell-shape)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plain Beta</td>
<td>0.031</td>
<td>0.028</td>
<td>0.029</td>
</tr>
<tr>
<td>Macro-Beta</td>
<td>0.031</td>
<td>0.025</td>
<td>0.027</td>
</tr>
<tr>
<td>Micro-Beta</td>
<td>0.031</td>
<td>0.028</td>
<td>0.029</td>
</tr>
</tbody>
</table>
FIGURE 7: Exact and Estimated Right Tail, \( \alpha = .5, \beta = .5 \), \( n = 100 \)
True: Solid, Plain Beta: Dashes, Macro-Normalized Beta: Short Dashes,
Micro-Normalized Beta: Dots and Dashes.
The micro-beta method dominates the macro-beta method, which in turn dominates the standard beta-kernel.

The magnitude of the average m.s.e is much larger than in Renault, Scaillet (2004).

Indeed, they use a Gauss-Legendre quadrature approach to compute the integrated squared errors, which creates a spurious smoothing effect.
The standard Gaussian kernel estimator:

$$\hat{f}_n^3(x) = 1_{[0,1]}(x) \frac{1}{nh} \sum_{i=1}^{n} \varphi \left( \frac{x - X_i}{h} \right)$$

The truncated Gaussian kernel estimator:

$$\hat{f}_n^4(x) = 1_{[0,1]}(x) \frac{1}{nh} \sum_{i=1}^{n} \frac{\varphi \left( \frac{x - X_i}{h} \right)}{\Phi \left( \frac{1 - X_i}{h} \right) - \Phi \left( \frac{-X_i}{h} \right)};$$

Apply a Gaussian kernel to the data preliminary transformed by the logistic transformation:

$$x \rightarrow \log(x/(1 - x))$$

$$\hat{f}_n^5(x) = \frac{1_{[0,1]}(x)}{nhx(1 - x)} \sum_{i=1}^{n} \varphi \left[ \frac{\log \left( \frac{x}{1 - x} \right) - \log \left( \frac{X_i}{1 - X_i} \right)}{h} \right]$$

C. GOURIEROUX, A. MONFORT

(Non) CONSISTENCY OF THE BETA KERNEL ESTIMATOR FOR RECOVERY RATE DISTRIBUTION
Table 2: Comparison of the Micro-beta Method with Gaussian Methods

<table>
<thead>
<tr>
<th>α = 0.5, β = 0.5 (U-shape)</th>
<th>Global</th>
<th>Left Tail</th>
<th>Right Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain Beta</td>
<td>0.279</td>
<td>1.334</td>
<td>1.324</td>
</tr>
<tr>
<td>Micro-Beta</td>
<td>0.135</td>
<td>0.611</td>
<td>0.615</td>
</tr>
<tr>
<td>Gaussian</td>
<td>0.598</td>
<td>2.936</td>
<td>2.939</td>
</tr>
<tr>
<td>Truncated Gaussian</td>
<td>0.493</td>
<td>2.219</td>
<td>2.225</td>
</tr>
<tr>
<td>Transformed Gaussian</td>
<td>0.151</td>
<td>0.739</td>
<td>0.687</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>α = 2.5, β = 0.5 (Skewed shape)</th>
<th>Global</th>
<th>Left Tail</th>
<th>Right Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain Beta</td>
<td>0.687</td>
<td>0.0013</td>
<td>6.71</td>
</tr>
<tr>
<td>Micro-Beta</td>
<td>0.217</td>
<td>0.0013</td>
<td>2.02</td>
</tr>
<tr>
<td>Gaussian</td>
<td>1.729</td>
<td>0.0006</td>
<td>17.05</td>
</tr>
<tr>
<td>Truncated Gaussian</td>
<td>1.378</td>
<td>0.0008</td>
<td>13.06</td>
</tr>
<tr>
<td>Transformed Gaussian</td>
<td>0.226</td>
<td>0.0046</td>
<td>2.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>α = 2.5, β = 2.5 (Bell-shape)</th>
<th>Global</th>
<th>Left Tail</th>
<th>Right Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain Beta</td>
<td>0.031</td>
<td>0.028</td>
<td>0.029</td>
</tr>
<tr>
<td>Micro-Beta</td>
<td>0.031</td>
<td>0.028</td>
<td>0.027</td>
</tr>
<tr>
<td>Gaussian</td>
<td>0.027</td>
<td>0.016</td>
<td>0.016</td>
</tr>
<tr>
<td>Truncated Gaussian</td>
<td>0.029</td>
<td>0.024</td>
<td>0.024</td>
</tr>
<tr>
<td>Transformed Gaussian</td>
<td>0.032</td>
<td>0.027</td>
<td>0.026</td>
</tr>
</tbody>
</table>

The micro-beta kernel and the transformed Gaussian kernel perform well in all situations.
INTRODUCTION
THE NONCONSISTENCY OF THE BETA KERNEL APPROACH
IMPROVED BETA KERNEL METHODS
COMPARISON OF THE MICRO-BETA METHOD WITH ALTERNATIVES
CONCLUSION

C. GOURIEROUX, A. MONFORT

(Non) CONSISTENCY OF THE BETA KERNEL ESTIMATOR FOR

FIGURE 11: Exact and Estimated Right Tail, alpha = 0.5, beta = 0.5, n = 100
True: Solid, Plain Beta: Dashes, Transformed Gaussian: Short Dashes,
Micro-Normalized Beta: Dots and Dashes.
FIGURE 12: Exact and Estimated Right Tail, \( \alpha = 2.5, \beta = 0.5, n = 100 \)
- True: Solid
- Plain Beta: Dashes
- Transformed Gaussian: Short Dashes
- Micro-Normalized Beta: Dots and Dashes
FIGURE 13: Exact and Estimated Right Tail, $\alpha = .5$, $\beta = .5$, $n = 1000$
True: Solid, Plain Beta: Dashes, Transformed Gaussian: Short Dashes,
Micro-Normalized Beta: Dots and Dashes
INTRODUCTION
THE NONCONSISTENCY OF THE BETA KERNEL APPROACH
IMPROVED BETA KERNEL METHODS
COMPARISON OF THE MICRO-BETA METHOD WITH ALTERNATIVE APPROACHES
CONCLUSION

FIGURE 14: Exact and Estimated Right Tail, $\alpha=2.5, \beta=.5, n=1000$
True: Solid, Plain Beta: Dashes, Transformed Gaussian: Short Dashes,
Micro-Normalized Beta: Dots and Dashes

C. GOURIEROUX, A. MONFORT

(Non) CONSISTENCY OF THE BETA KERNEL ESTIMATOR FOR
To avoid important bias when computing a Credit VaR, both parametric and nonparametric estimation approaches have to account for the point mass at 0% (and 100%).

The plain beta kernel method is not the most appropriate to estimate the continuous part of the LGD distribution.