Supplementary Note S5. Computational implementation and reproducibility of the BLLU applications

This note complements the theoretical and estimation material in Supplementary Notes S1–S4. It collects the numerical routines used for evaluating the BLL/BLLU probability mass functions, checking limiting cases, and reproducing the empirical goodness-of-fit (GOF) analyses. The accompanying reproducibility folder contains the R scripts and the original application data.

S5.1. Numerical evaluation of the PMF and distributional characteristics

Script 01_BLL_PMF_and_moments.R evaluates the probability mass function by integrating the binomial kernel against the two-piece LLUD mixing density. The integration interval is split at the density breakpoint. The script then checks that the computed probabilities sum to one and calculates the corresponding mean and variance from the finite support.

Script 02_BLLU_limit_cases_and_checks.R evaluates the BLLU PMF and CDF through the LLU quantile representation. Numerical integration is split at the quantile breakpoint $u = 1/2$. The script illustrates the large-n and boundary-parameter regimes and includes normalization and CDF endpoint checks.

The supplied routines use numerical integration and midpoint quadrature. No analytical derivatives of the PMF are used in these calculations. The parameter-estimation routines and their optimization settings are documented separately in Supplementary Notes S3 and S4.

S5.2. Empirical goodness-of-fit tests and Monte Carlo calibration

Script 07_Empirical_GOF_and_MC_calibration.R reproduces the discrete Cramér–von Mises (CvM) and Anderson–Darling (AD) statistics for the Binomial, Beta-Binomial, Kumaraswamy-Binomial, and BLLU models fitted to the SO_2 and traffic-accident datasets. The Excel workbook BLLU_Application_Data.xlsx contains the raw observations and grouped frequency tables. The script verifies that the grouped frequencies agree with the raw observations before performing the GOF calculations.

For the BLLU model, the PMF is evaluated using the LLU quantile representation and midpoint quadrature with $K = 2{,}000$ points. The Monte Carlo calibration uses $B = 50{,}000$ multinomial samples for each model–dataset combination, generated under the fitted null distribution. The fitted parameter values are held fixed across the replications. The reported Monte Carlo $p-$values use the standard +1 correction, and the corresponding Monte Carlo standard errors are calculated.

S5.3. Reproducibility details

Component	Setting
Random-number seed	123 (GOF Monte Carlo calibration)
Monte Carlo replications	$B = 50{,}000$ per model–dataset combination
BLLU midpoint quadrature	$K = 2{,}000$
Required non-base R package	readxl (for importing the Excel workbook)
Estimation and optimization	See Supplementary Notes S3 and S4

The GOF script prints the R session information at completion, including the R version and installed package versions, to facilitate reporting of the exact computational environment. The complete execution instructions are provided in the accompanying README file.

Files supplied

The accompanying folder contains the three R scripts listed above, the original Excel workbook, and a README file describing their execution. Estimation and Monte Carlo simulation code for MoM, MLE, and quantile procedures remains in the existing supplementary materials, avoiding duplication of the material already documented in Supplementary Notes S3 and S4.