Supplementary Material

1 Final Parameter Updates

1. Sampling $\boldsymbol{\pi}$

$$\pi_i | \underset{\sim}{t} \sim \text{Dirichlet}(\underset{\sim}{n_i} + \underset{\sim}{t}), \qquad i = 1, 2, \ldots, d$$

   where, $n_i = (n_{i1}, n_{i2}, \ldots, n_{id})$

2. Sampling $\underset{\sim}{u}$

$$P(\underset{\sim}{u}|t) \propto \prod_{i=1}^{d} e^{-u_i} u_i^{\sum_{j=1}^{d} t_j - 1} \mathbb{1}_{[u_i > 0]}$$

   That is,

$$u_j | t \overset{ind}{\sim} \text{Gamma}(t, 1) \text{ for all } j = 1, \ldots, d$$

3. Sampling $\underset{\sim}{w}$

$$P(\underset{\sim}{w}|\alpha, \underset{\sim}{t}) \propto \prod_{j=1}^{d} e^{-(t_j + t_{j+1} + \cdots + t_d) w_j} w_j^{\alpha - 1} \mathbb{1}_{[w_j > 0]}$$

   That is,

$$w_j | \alpha, \underset{\sim}{t} \overset{ind}{\sim} \text{Gamma}(\alpha, \sum_{k=j}^{d} t_k) \text{ for all } j = 1, \ldots, d$$

4. Sampling $\underset{\sim}{t}$

$$P(\underset{\sim}{t}|\boldsymbol{\pi}, \underset{\sim}{u}, \underset{\sim}{w}) \propto \prod_{j=1}^{d} f_j(t_j)$$

   where,

$$f_j(t) \propto \frac{1}{\Gamma(t)^d} e^{-(b_0 + \sum_{k=1}^{j} w_k - \sum_{i=1}^{d} \log \pi_{ij} - \sum_{i=1}^{d} \log u_i) t} t^{\delta_j - 1} \mathbb{1}_{[t > 0]}, \qquad \delta_j = \begin{cases} \alpha \text{ if } j = 1, 2, \cdots, d-1 \\ \beta \text{ if } j = d \end{cases}$$

That is, f_j is a tilted Gamma density with parameters $d, \delta_j, B_j = b_0 + \sum_{k=1}^{j} w_k - \sum_{i=1}^{d} \log \pi_{ij} - \sum_{i=1}^{d} \log u_i$ and $\alpha_0 = t$.

2 Proofs of Auxiliary Results

2.1 Proof of Lemma 1

Proof Let $\underset{\sim}{Y} = (Y_1, Y_2, \ldots, Y_k)$ be a random vector such that the components are independently distributed according to

$$Y_i \sim \text{Gamma}(\alpha_i), \qquad i = 1, 2, \ldots, k.$$

Define the normalized vector $\underset{\sim}{X} = (X_1, X_2, \ldots, X_k)$ as

$$X_i = \frac{Y_i}{\sum_{j=1}^{k} Y_j}, \qquad i = 1, 2, \ldots, k.$$

Then,

$$\underset{\sim}{X} \sim \text{Dirichlet}(\alpha_1, \alpha_2, \ldots, \alpha_k).$$

Similarly,

$$\frac{1}{\sum_{i=2}^{k} Y_i}(Y_2, Y_3, \ldots, Y_k) \sim \text{Dirichlet}(\alpha_2, \alpha_3, \ldots, \alpha_k).$$

Now,

$$\begin{aligned}
\frac{1}{\sum_{i=2}^{k} Y_i}(Y_2, Y_3, \ldots, Y_k) &= \frac{1}{\sum_{i=1}^{k} Y_i - Y_1}(Y_2, Y_3, \ldots, Y_k) \\
&= \frac{1}{1 - \frac{Y_1}{\sum_{i=1}^{k} Y_i}} \left(\frac{Y_2}{\sum_{i=1}^{k} Y_i}, \frac{Y_3}{\sum_{i=1}^{k} Y_i}, \ldots, \frac{Y_k}{\sum_{i=1}^{k} Y_i} \right) \\
&= \frac{1}{1 - X_1}(X_2, X_3, \ldots, X_k),
\end{aligned}$$

which implies,

$$\frac{1}{1 - X_1}(X_2, X_3, \ldots, X_k) \sim \text{Dirichlet}(\alpha_2, \alpha_3, \ldots, \alpha_k).$$

□

2.2 Proof of Theorem 2

Proof Let $(A_1 = I_1, A_2 = I_2, \ldots, A_r = I_r)$ be a finite partition $\Theta = \mathbb{N}$, the set of positive integers. Then, using the definition of Dirichlet process, we get,

$$\begin{aligned}
&\big(G_j(A_1), G_j(A_2), \ldots, G_j(A_r)\big) \sim \text{Dirichlet}\left(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_r)\right) \\
&\implies \left(\sum_{j \in I_1} \pi_{ij}, \ldots, \sum_{j \in I_r} \pi_{ij} \right) \sim \text{Dirichlet}\left(\alpha_0 \sum_{j \in I_1} \beta_j, \ldots, \alpha_0 \sum_{j \in I_r} \beta_j \right) \qquad (1)
\end{aligned}$$

We make a partition $(\{1, 2, \ldots, j-1\}, \{j\}, \{j+1, j+2 \ldots\})$ of $\mathbb{N}$ such that,

$$\left(\sum_{k=1}^{j-1} \pi_{ik}, \pi_{ij}, \sum_{k=j+1}^{\infty} \pi_{ik} \right) \sim \text{Dirichlet}\left(\alpha_0 \sum_{k=1}^{j-1} \gamma_k, \alpha_0 \gamma_j, \alpha_0 \sum_{k=j+1}^{\infty} \gamma_k \right)$$

Then by removing the first element and using Lemma 2.1,

$$\frac{1}{1-\sum_{k=1}^{j-1}\pi_{ik}}\left(\pi_{ij}, \sum_{k=j+1}^{\infty}\pi_{ik}\right) \sim \text{Dirichlet}\left(\alpha_0\gamma_j, \alpha_0\sum_{k=j+1}^{\infty}\gamma_k\right) \tag{2}$$

$\pi_{ij} = \pi'_{ij}\prod_{k=1}^{j-1}(1-\pi'_{ik})$ implies that

$$\pi'_{ij} = \frac{\pi_{ij}}{1-\sum_{k=1}^{j-1}\pi_{ik}}.$$

Again,

$$\frac{\sum_{k=j+1}^{\infty}\pi_{ik}}{1-\sum_{k=1}^{j-1}\pi_{ik}} = \frac{1-\sum_{k=1}^{j}\pi_{ik}}{1-\sum_{k=1}^{j-1}\pi_{ik}} = 1-\pi'_{ij},$$

and

$$\sum_{k=j+1}^{\infty}\gamma_k = 1-\sum_{k=1}^{j}\gamma_k.$$

Hence, from (2),

$$(\pi'_{ij}, 1-\pi'_{ij}) \sim \text{Dirichlet}\left(\alpha_0\gamma_j, \alpha_0\sum_{k=1}^{j}(1-\gamma_k)\right).$$

$$\implies \pi'_{ij} \sim \text{ Beta }(\alpha_0\gamma_j, \alpha_0(1-\sum_{k=1}^{j}\gamma_k)) \qquad i,j=1,2,\ldots$$

□

2.3 Proof of Lemma 3

Proof: Given that each ν_j is independently and identically distributed as a Beta(α,β) and the random vector $\underset{\sim}{\gamma} = (\gamma_1,\gamma_2,\ldots,\gamma_d)$ is defined as

$$\gamma_j = \nu_j\prod_{k=1}^{j-1}(1-\nu_k)$$

for all $j=1,2,\ldots,d-1$ and

$$\gamma_d = \prod_{k=1}^{d-1}(1-\nu_k)$$

such that $\sum_{j=1}^{d}\gamma_j = 1$. Then it follows from the construction of Generalized Dirichlet Distribution that

$$\underset{\sim}{\gamma}_{d-1} = (\gamma_1,\gamma_2,\ldots,\gamma_{d-1}) \sim \text{GD}_{d-1}(\underbrace{\alpha,\alpha,\ldots,\alpha}_{\text{d-1 times}};\underbrace{\beta,\beta,\ldots,\beta}_{\text{d-1 times}})$$

Similarly, given that for each i, the random variables π'_{ij}'s are independently distributed as Beta$(\alpha_0\gamma_j, \alpha_0(1-\sum_{k=1}^{j}\gamma_k))$ with

$$\pi_{ij} = \pi'_{ij}\prod_{k=1}^{j-1}(1-\pi'_{ik})$$

for all $j = 1, 2, \ldots, d-1$ and

$$\pi_{id} = \prod_{k=1}^{d-1} (1 - \pi'_{ik})$$

such that $\sum_{j=1}^{d} \pi_{ij} = 1$. Then again, by the construction of Generalized Dirichlet distribution, for all $i = 1, 2, \ldots, d$, we have

$$\pi_{i1}, \ldots, \pi_{id-1} \sim \text{GD}\left(\alpha_0\gamma_1, \ldots, \alpha_0\gamma_{d-1}; \alpha_0(1-\gamma_1), \ldots, \alpha_0(1 - \sum_{k=1}^{d-1}\gamma_k)\right).$$

Hence the probability density function of $(\pi_{i1}, \ldots, \pi_{id-1})$ becomes

$$f(\pi_{i1}, \ldots, \pi_{id-1}) \propto \pi_{id-1}^{\alpha_0\gamma_{d-1}} (1 - \pi_{i1} - \cdots - \pi_{id-1})^{\alpha_0(1-\sum_{k=1}^{d-1}\gamma_k)}$$
$$\prod_{j=1}^{d-2} \pi_{ij}^{\alpha_0\gamma_j} (1 - \pi_{i1} - \cdots - \pi_{ij})^{\alpha_0(1-\sum_{k=1}^{j}\gamma_k) - \alpha_0\gamma_{j+1} - \sum_{k=1}^{j+1}\gamma_k}.$$

Now, the denominator corresponding to the term $1 - \pi_{i1} - \cdots - \pi_{ij}$ becomes 0 for all $j = 1, \ldots, d-2$. Moreover, given that $1 - \pi_{i1} - \cdots - \pi_{id-1} = \pi_{id}$ and $1 - \sum_{k=1}^{d-1}\gamma_k = \gamma_d$, the probability density function of $\underset{\sim}{\pi_i} = (1 - \pi_{i1} - \cdots - \pi_{id})$ becomes

$$f(\pi_{i1}, \ldots, \pi_{id}) \propto \pi_{i,d-1}^{\alpha_0\gamma_{d-1}} \pi_{id}^{\alpha_0\gamma_d} \prod_{j=1}^{d-2} \pi_{ij}^{\alpha_0\gamma_j},$$

which implies,

$$f(\pi_{i1}, \ldots, \pi_{id}) \propto \prod_{j=1}^{d} \pi_{ij}^{\alpha_0\gamma_j} \quad 0 < \pi_{ij} < 1 \text{ and, } \sum_{j=1}^{d} \pi_{ij} = 1 \quad \forall\, i = 1, \ldots, d.$$

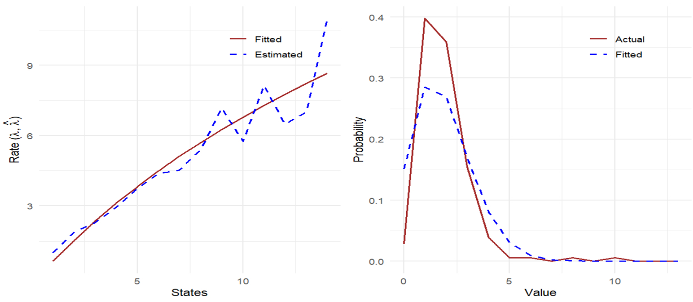
This is the probability density function of a Dirichlet$(\alpha_0 \underset{\sim}{\gamma})$ distribution. □

3 Rationale for Choices of Simulated Datasets

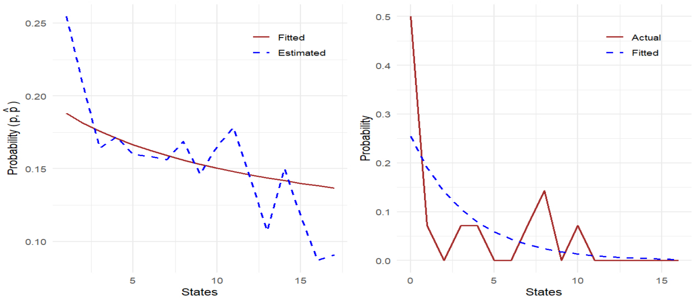
To motivate the choice of simulation distributions, we focus on two widely used discrete families with countable support: the Geometric and the Poisson distributions. Empirical validation is performed using the two real-life datasets analyzed in the main paper. For the USO stock volume data, we fit a Poisson distribution to a representative row of the empirical TPM, employing the same rate specification $\lambda_i = \log(i + c)$ as in our simulation design. For the Heathrow rainfall data, we fit a Geometric distribution to a representative row of the empirical TPM, using the functional form $p_i = \lambda_i^{-1}$. The respective distributions were chosen because they offered the best empirical fit to the observed data. The corresponding graphs have been provided in the next section.

The constant c must satisfy $c > e-1$ to ensure that the values of p_i remain bounded strictly between zero and unity. In practice, we experimented with several choices of $c \geq 2$ and evaluated the resulting fits both numerically and visually. The coefficient of determination (R^2), as a measure of goodness of fit, together with graphical inspection, indicated that values of c in the interval $5 \leq c \leq 20$ consistently yield strong performance across both cases, with only minor variations in R^2 and close graphical conformity of the fitted curves. In all the subsequent analyses, we fix $c = 9$, as it provides an excellent fit with high R^2 and strong visual agreement in both cases. Alternative choices of c within this interval are observed to yield similar results.

4 Additional Tables and Graphs



(a) Poisson



(b) Geometric

Fig. S1 Empirical validation of distributional choices for the simulation study. Panel (a) shows the Poisson fit on a representative row of the USO TPM and the comparison between $\hat{\lambda}_i$ and its functional form. Panel (b) shows the Geometric fit on a representative row of the Heathrow TPM and the comparison between $\hat{p}_i$ and its functional form. In both cases, the parameter functions and the distributions align well with the observed data, supporting their use in the simulation study.

(a) Heathrow Rain

(b) USO Stock Volume

Fig. S2 Actual vs Predicted Series using the three methods (Training:Test = 50:50) - (a) GHSBP: Bayesian Estimation with generalized hierarchical stick-breaking prior, (b) HSBP: Bayesian estimation with hierarchical stick-breaking prior and (c) MLE: maximum likelihood estimation

Table S1 MAE (%) and RMSE (%) for GHSBP under Poisson distribution. Bold values indicate the lowest RMSE. Each block varies one hyperparameter, keeping the others fixed.

	α	β	b_0	MAE	RMSE
Vary b_0	0.5	2	0.05	0.981	2.425
	0.5	2	0.1	0.981	2.422
	0.5	2	0.2	0.971	**2.419**
	0.5	2	0.25	0.979	2.428
	0.5	2	0.5	0.983	2.443
	0.5	2	1	0.981	2.428
Vary α	0.45	2	0.1	0.987	2.426
	0.5	2	0.1	0.981	2.422
	1	2	0.1	0.995	2.459
	2	2	0.1	0.994	2.475
	0.4	2	0.2	0.984	2.427
	0.5	2	0.2	0.971	**2.419**
	1	2	0.2	0.988	2.440
	0.7	0.5	1	0.987	2.448
	1	0.5	1	0.991	2.449
	2	1	10	1.026	2.564
	1	1	10	1.016	2.524
Vary β	0.5	1	0.1	0.987	2.436
	0.5	2	0.1	0.981	2.422
	0.5	3	0.1	0.983	2.442
	0.5	1	0.2	0.981	2.438
	0.5	2	0.2	0.971	**2.419**
	0.5	3	0.2	0.981	2.430

Table S2 MAE (%) and RMSE (%) for GHSBP under Geometric distribution. Bold values indicate the lowest RMSE. Each block varies one hyperparameter, keeping the others fixed.

	α	β	b_0	MAE	RMSE
Vary b_0	0.5	2	0.05	1.013	1.961
	0.5	2	0.1	1.018	**1.956**
	0.5	2	0.2	1.016	1.968
	0.5	2	0.25	1.026	1.981
	0.5	2	0.5	1.031	1.996
	0.5	2	1	1.036	2.022
Vary α	0.45	2	0.1	1.031	1.990
	0.5	2	0.1	1.018	**1.956**
	1	2	0.1	1.000	2.081
	2	2	0.1	1.014	2.224
	0.4	2	0.2	1.042	1.987
	0.5	2	0.2	1.016	1.968
	1	2	0.2	0.996	2.065
	0.7	0.5	1	1.023	2.040
	1	0.5	1	1.011	2.069
	2	1	10	1.061	2.349
	1	1	10	1.092	2.336
Vary β	0.5	1	0.1	1.026	1.985
	0.5	2	0.1	1.018	**1.956**
	0.5	3	0.1	1.027	1.992
	0.5	1	0.2	1.022	1.985
	0.5	2	0.2	1.016	1.968
	0.5	3	0.2	1.032	2.013