BAYESIAN STOCHASTIC FRONTIER MODELS UNDER THE SKEW-NORMAL HALF-NORMAL SETTINGS WITH PYMC

Yejin Hwang **Zak Zapata** **Emily Liu**
Texas A&M University-Corpus Christi
6300 Ocean Dr, Corpus Christi, TX 78412
yhwang@islander.tamucc.edu zzapata2@islander.tamucc.edu cliu7@islander.tamucc.edu

May 13, 2025

ABSTRACT

This paper implements a Bayesian stochastic frontier model under the skew-normal and half-normal error specifications using PyMC in Python. Following the methodology from Wei et al. (2025), we replicate their simulation and real-data analysis using a custom Bayesian MCMC workflow. We emphasize posterior analysis, posterior predictive checks, and model comparison in order to evaluate the effectiveness of the model proposed by Wei et al (2025).

Keywords Bayesian statistics · Efficiency · Markov chain Monte Carlo · Skew-normal distribution · Stochastic frontier model · PyMC

1 Introduction

Stochastic Frontier Analysis (SFA) has become a widely used framework for estimating production efficiency in economics and operations research. Originally introduced by Aigner et al. (1977), the SFA framework decomposes deviations from the production frontier into two components: random noise and inefficiency. This approach provides a statistically rigorous basis for identifying firms or units that underperform relative to an estimated best-practice frontier.

The classical stochastic frontier model proposed by Aigner et al. (1977) defines the composed error term as the sum of a symmetric noise component, modeled by a normal distribution, and a one-sided inefficiency term, modeled by a half-normal distribution. While convenient, these assumptions can be overly restrictive and may fail to capture the asymmetric nature of real-world inefficiency. To address this, more flexible error specifications have been proposed, such as the skew-normal distribution for noise components, as discussed in Wei et al. (2025). These extensions enable the model to better account for asymmetry in the data, improving both fit and interpretability.

While MLE has been widely used in stochastic frontier models, it often struggles to provide robust inference under asymmetric error structures, particularly for the skewness parameter λ (Wei et al., 2025). In contrast, Bayesian approaches using MCMC sampling offer a reliable alternative by approximating the full posterior distribution and mitigating convergence issues.

Stochastic Frontier Analysis (SFA) has been widely implemented in R through packages such as `frontier`, `sfaR`, and `brms`, and has been extensively applied in domains such as healthcare

(Worthington (2022)) and higher education efficiency research (Johnes and Portela (2019)). In contrast, while PyMC is a widely used framework for Bayesian inference,dedicated implementations or case studies involving SFA models remain limited in its official documentation and academic usage (Salvatier et al. (2016)).

Our contributions are threefold: (1) we replicate the SN-HN stochastic frontier model in PyMC, providing an accessible and transparent implementation for Python users; (2) we apply the model to simulated production data generated from a SN-HN distribution and evaluate the model's performance using posterior analysis; and (3)we compare the performance of proposed by (Wei et al., 2025) which employs a skew-normal likelihood with that of a conventional model based on a normal likelihood to assess the impact of distributional asymmetry on posterior inference.

2 Preliminaries

2.1 Normal and Skew-Normal Distributions

The normal (Gaussian) distribution is a symmetric, bell-shaped distribution widely used in statistical modeling due to its analytical tractability. A random variable X following a normal distribution with mean μ and variance σ^2 is denoted as $X \sim \mathcal{N}(\mu, \sigma^2)$, and its probability density function is:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

While useful, the normal distribution assumes symmetry about the mean, which may not be appropriate when modeling real-world data that exhibit skewness.

To address this limitation, the skew-normal distribution introduces an additional shape parameter λ to model asymmetry. A skew-normal distributed variable $X \sim SN(\mu, \sigma^2, \lambda)$ has the following density:

$$f_Y(y) = \frac{2}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\frac{\lambda(y-\mu)}{\sigma}\right)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. The sign and magnitude of λ control the direction and degree of skewness. When $\lambda = 0$, the skew-normal distribution reduces to the standard normal distribution.

2.2 Stochastic Frontier Models

Stochastic Frontier Analysis (SFA) is a methodology used to measure the efficiency of decision-making units (e.g., firms, hospitals, or regions) relative to a theoretical production frontier. Introduced by Aigner et al. (1977), the canonical model is given by:

$$Y_i = f(X_i; \beta) + V_i - U_i$$

Here, Y_i is the logarithm of the observed output, $f(X_i; \beta)$ is a deterministic frontier function parameterized by β, V_i is a symmetric random noise term (often assumed normal), and U_i is a non-negative inefficiency term.

In traditional SFA, $V_i \sim N(0, \sigma_v^2)$ and $U_i \sim HN(0, \sigma_u^2)$, where the half-normal distribution ensures $U_i \geq 0$. The one-sided nature of U_i reflects the fact that inefficiency can only reduce output, not increase it.

However, recent extensions such as those proposed by Wei et al. (2025) allow for V_i to follow a skew-normal distribution, improving model flexibility in cases where measurement noise is not symmetric. This allows for more accurate inference in settings where classical assumptions may be violated.

3 Methodology

3.1 Model Specification and Simulation Design

To evaluate the Bayesian stochastic frontier model under the skew-normal and half-normal setting, we conduct a simulation study following the structure proposed in Wei et al. (2025). Our primary objective is to investigate the predictive performance of the model for varying values of the skewness parameter λ, with an emphasis placed on comparing the performance of the proposed model to predict inefficiency compared to a model using the likelihood of the normal distribution.

Model and Data Generation

We generate synthetic data according to the following stochastic production function.

$$Y = \alpha + \beta X + V - U$$

where Y represents the logarithm of the output variable, X denotes the logarithm of a single explanatory variable, α is the intercept, β is the regression coefficient, U represents the inefficiency error term, V is the measurement error term and ξ is the location parameter.

$$\begin{aligned} X &\sim \mathrm{N}(1,1) \\ U &\sim \mathrm{HN}(0,\sigma_u^2) \\ Y &\sim \mathrm{SN}(\xi,\sigma_v^2,\lambda) \\ \xi &= \alpha + \beta X - U \end{aligned}$$

The true parameter values used in simulation are:

$$\alpha = 5, \quad \beta = 2, \quad \sigma_u = 1, \quad \sigma_v = 1.5$$

Prior Distributions

For the Bayesian analysis, we adopt the following prior distributions:

$$\begin{aligned} \alpha &\sim \mathrm{N}(\mu_\alpha,\sigma_\alpha^2) \\ \beta &\sim \mathrm{N}(\mu_\beta,\sigma_\beta^2) \\ \sigma_v^2 &\sim \mathrm{IG}(\alpha_v,\beta_v) \\ \sigma_u^2 &\sim \mathrm{IG}(\alpha_u,\beta_u) \\ \lambda &\sim \mathrm{TN}(\mu_\lambda,\sigma_\lambda^2;a,b) \end{aligned}$$

To assess the robustness of posterior inference for λ, we simulate data across the following scenarios:

- **Skewness values:** $\lambda \in \{-0.5, -1, -1.5, -2, -5\}$
- **Sample sizes:** $n \in \{50, 100, 200, 500\}$

This results in a total of 20 simulation scenarios. For each scenario, n synthetic observations are generated and used as input for Bayesian inference.

Inference Procedure

Posterior inference is conducted using the No-U-Turn Sampler (NUTS) provided by the PyMC library. For each simulation scenario, we run four independent chains, each with 6000 iterations including 2000 warm-up steps. Convergence is assessed using the Gelman-Rubin statistic ($\hat{R}$) and effective sample size (ESS), as computed by the ArviZ package.

4 Result

4.1 Posterior summaries

To assess how well each model recovers the underlying parameters, we summarize the posterior estimates for all scenarios. We present the results under both the Skew-Normal and Normal assumptions across different sample sizes and skewness levels (λ). These tables help to visualize how the flexibility of the model impacts the quality of the inference.

In addition, trace plots were generated to analyze the MCMC convergence for all posterior distributions of parameters. We note that for almost all sample sizes and λ values, MCMC convergence was achieved. This indicates reliability in our generated model's posterior distribution samples.

Table 1: Posterior summaries under Skewness (left) and Normal (right) models for $\lambda = -0.5$

	Skewness Model					**Normal Model**			
Sample	α	β	σ_v	σ_u	λ	α	β	σ_v	σ_u
True Value	5	2	1.5	1	-0.5	5	2	1.5	1
n = 50									
Mean	4.366	2.325	1.629	0.815	-0.663	3.711	2.329	1.435	0.823
Sd	0.550	0.186	0.275	0.352	0.456	0.396	0.183	0.177	0.346
3 HDI	3.358	1.993	1.148	0.267	-1.455	2.988	1.985	1.104	0.287
97 HDI	5.404	2.690	2.162	1.477	0.000	4.476	2.673	1.773	1.480
ESS(bulk)	5465.000	23718.000	8645.000	2414.000	10329.000	3687.000	22980.000	7634.000	2336.000
ESS(tail)	10265.000	17303.000	9657.000	4585.000	10462.000	5999.000	17068.000	8261.000	4490.000
R-hat	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n = 100									
Mean	5.570	1.906	1.466	0.913	-0.890	4.905	1.885	1.235	0.931
Sd	0.463	0.122	0.264	0.397	0.593	0.351	0.122	0.162	0.399
3 HDI	4.662	1.676	0.983	0.313	-1.908	4.290	1.658	0.917	0.303
97 HDI	6.375	2.136	1.978	1.672	0.000	5.576	2.114	1.536	1.674
ESS(bulk)	1963.000	11191.000	1390.000	628.000	9070.000	540.000	10727.000	645.000	392.000
ESS(tail)	5265.000	15358.000	1110.000	1082.000	10262.000	1366.000	14890.000	1019.000	1079.000
R-hat	1.000	1.000	1.000	1.010	1.000	1.010	1.000	1.010	1.020
n = 200									
Mean	5.251	2.073	1.651	0.949	-1.017	4.526	2.074	1.297	1.135
Sd	0.433	0.100	0.267	0.386	0.549	0.382	0.101	0.162	0.442
3 HDI	4.413	1.882	1.149	0.351	-1.877	3.852	1.880	0.986	0.350
97 HDI	6.039	2.258	2.123	1.683	0.000	5.216	2.260	1.570	1.866
ESS(bulk)	2368.000	20219.000	824.000	401.000	4850.000	282.000	15763.000	316.000	258.000
ESS(tail)	4975.000	16888.000	1326.000	939.000	8651.000	1109.000	17866.000	933.000	981.000
R-hat	1.000	1.000	1.010	1.010	1.000	1.010	1.000	1.010	1.010
n = 500									
Mean	5.004	1.916	1.516	0.746	-0.519	4.489	1.917	1.388	0.761
Sd	0.355	0.065	0.141	0.257	0.326	0.233	0.066	0.074	0.268
3 HDI	4.337	1.793	1.271	0.310	-1.062	4.091	1.795	1.246	0.301
97 HDI	5.632	2.038	1.795	1.227	0.000	4.939	2.041	1.525	1.248
ESS(bulk)	2131.000	35984.000	1878.000	491.000	11070.000	478.000	41664.000	702.000	410.000
ESS(tail)	4773.000	18542.000	2327.000	994.000	12011.000	1386.000	18265.000	1380.000	1112.000
R-hat	1.000	1.000	1.000	1.010	1.000	1.010	1.000	1.010	1.010

Table 2: Posterior summaries under Skewness (left) and Normal (right) models for $\lambda = -1$

	Skewness Model					Normal Model			
Sample	α	β	σ_v	σ_u	λ	α	β	σ_v	σ_u
True Value	5	2	1.5	1	-1	5	2	1.5	1
n = 50									
Mean	4.250	2.297	1.615	0.843	-0.873	3.462	2.304	1.362	0.831
Sd	0.551	0.176	0.294	0.361	0.548	0.392	0.176	0.177	0.353
3 HDI	3.250	1.958	1.095	0.280	-1.813	2.752	1.982	1.032	0.277
97 HDI	5.297	2.628	2.177	1.522	-0.001	4.212	2.644	1.699	1.488
ESS(bulk)	5670.000	24707.000	6714.000	2062.000	11242.000	3518.000	19627.000	6012.000	2411.000
ESS(tail)	9798.000	17158.000	6407.000	3675.000	10937.000	4575.000	16648.000	5388.000	3684.000
R-hat	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n = 100									
Mean	5.260	1.944	1.433	0.884	-1.247	4.500	1.921	1.118	0.932
Sd	0.409	0.110	0.269	0.372	0.693	0.336	0.114	0.160	0.388
3 HDI	4.472	1.743	0.935	0.291	-2.378	3.888	1.698	0.793	0.306
97 HDI	6.000	2.157	1.946	1.578	0.000	5.125	2.128	1.397	1.649
ESS(bulk)	2044.000	13120.000	1475.000	656.000	4312.000	648.000	12117.000	777.000	543.000
ESS(tail)	4675.000	18265.000	1933.000	1740.000	5353.000	1542.000	16429.000	1131.000	1287.000
R-hat	1.000	1.000	1.000	1.000	1.000	1.010	1.000	1.010	1.010
n = 200									
Mean	4.996	2.048	1.518	0.955	-1.281	4.301	2.048	1.092	1.218
Sd	0.336	0.089	0.271	0.371	0.586	0.317	0.090	0.149	0.369
3 HDI	4.343	1.880	1.009	0.362	-2.177	3.688	1.884	0.829	0.497
97 HDI	5.615	2.217	1.980	1.636	-0.016	4.853	2.223	1.368	1.849
ESS(bulk)	2319.000	13976.000	554.000	346.000	1945.000	327.000	13632.000	349.000	296.000
ESS(tail)	4819.000	15236.000	1260.000	832.000	3236.000	823.000	16202.000	1405.000	591.000
R-hat	1.000	1.000	1.010	1.010	1.000	1.010	1.000	1.010	1.010
n = 500									
Mean	4.714	1.911	1.357	0.730	-0.635	4.193	1.910	1.205	0.761
Sd	0.323	0.058	0.143	0.251	0.361	0.223	0.058	0.076	0.263
3 HDI	4.114	1.802	1.100	0.301	-1.219	3.796	1.799	1.053	0.316
97 HDI	5.308	2.018	1.631	1.202	0.000	4.607	2.016	1.338	1.240
ESS(bulk)	1843.000	38217.000	1378.000	432.000	10567.000	457.000	33646.000	617.000	409.000
ESS(tail)	5258.000	18156.000	2124.000	1085.000	10653.000	1264.000	17482.000	1229.000	1099.000
R-hat	1.000	1.000	1.000	1.000	1.000	1.010	1.000	1.010	1.010

Table 3: Posterior summaries under Skewness (left) and Normal (right) models for $\lambda = -1.5$

	Skewness Model					Normal Model			
Sample	α	β	σ_v	σ_u	λ	α	β	σ_v	σ_u
True Value	5	2	1.5	1	-1.5	5	2	1.5	1
n = 50									
Mean	4.307	2.268	1.620	0.851	-1.212	3.384	2.280	1.272	0.861
Sd	0.522	0.165	0.313	0.373	0.664	0.392	0.165	0.179	0.373
3 HDI	3.323	1.950	1.059	0.281	-2.322	2.677	1.964	0.933	0.278
97 HDI	5.289	2.571	2.221	1.565	0.000	4.137	2.583	1.623	1.551
ESS(bulk)	5290.000	20233.000	5400.000	1997.000	9935.000	2905.000	21733.000	4225.000	2064.000
ESS(tail)	9661.000	18229.000	6300.000	3871.000	9283.000	4501.000	18216.000	4072.000	3901.000
R-hat	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n = 100									
Mean	5.130	1.971	1.414	0.893	-1.638	4.329	1.947	1.012	0.995
Sd	0.356	0.101	0.267	0.361	0.764	0.330	0.105	0.166	0.382
3 HDI	4.411	1.781	0.903	0.308	-2.983	3.740	1.756	0.686	0.349
97 HDI	5.750	2.160	1.895	1.557	-0.146	4.938	2.152	1.298	1.666
ESS(bulk)	1944.000	14719.000	1353.000	608.000	6242.000	391.000	12308.000	432.000	334.000
ESS(tail)	5782.000	16259.000	2069.000	1620.000	6239.000	2144.000	15869.000	1413.000	1637.000
R-hat	1.000	1.000	1.000	1.010	1.000	1.000	1.000	1.000	1.010
n = 200									
Mean	4.874	2.033	1.449	0.926	-1.587	4.208	2.034	0.941	1.275
Sd	0.277	0.081	0.266	0.363	0.614	0.261	0.081	0.134	0.304
3 HDI	4.361	1.883	0.910	0.327	-2.676	3.674	1.877	0.709	0.648
97 HDI	5.395	2.184	1.877	1.565	-0.316	4.664	2.185	1.206	1.799
ESS(bulk)	2058.000	13636.000	402.000	297.000	1367.000	373.000	12290.000	360.000	334.000
ESS(tail)	4929.000	16134.000	1070.000	994.000	2267.000	615.000	16443.000	738.000	523.000
R-hat	1.000	1.000	1.010	1.020	1.000	1.020	1.000	1.020	1.020
n = 500									
Mean	4.700	1.913	1.306	0.740	-0.951	4.090	1.911	1.061	0.827
Sd	0.286	0.053	0.166	0.247	0.438	0.223	0.053	0.086	0.270
3 HDI	4.136	1.812	0.998	0.294	-1.657	3.686	1.814	0.894	0.337
97 HDI	5.211	2.011	1.605	1.189	-0.055	4.491	2.013	1.205	1.297
ESS(bulk)	1816.000	36679.000	887.000	307.000	6643.000	243.000	27645.000	289.000	222.000
ESS(tail)	5575.000	18578.000	1726.000	538.000	11966.000	634.000	17353.000	849.000	464.000
R-hat	1.000	1.000	1.000	1.010	1.000	1.010	1.000	1.010	1.010

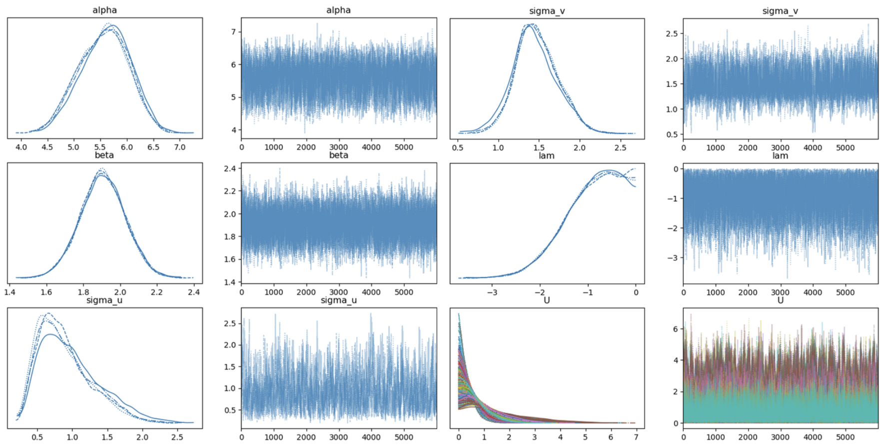
Table 4: Posterior summaries under Skewness (left) and Normal (right) models for $\lambda = -2$

	Skewness Model					Normal Model			
Sample	α	β	σ_v	σ_u	λ	α	β	σ_v	σ_u
True Value	5.000	2.000	1.500	1.000	-2.000	5.000	2.000	1.500	1.000
n = 50									
Mean	4.467	2.240	1.653	0.895	-1.689	3.380	2.261	1.204	0.902
Sd	0.479	0.156	0.327	0.391	0.772	0.400	0.158	0.186	0.392
3 HDI	3.524	1.942	1.050	0.288	-3.133	2.673	1.959	0.835	0.298
97 HDI	5.352	2.529	2.268	1.637	-0.239	4.157	2.552	1.546	1.650
ESS(bulk)	3956.000	20252.000	3826.000	1579.000	9154.000	2092.000	17246.000	2923.000	1606.000
ESS(tail)	7737.000	17800.000	4190.000	2832.000	8971.000	3225.000	17272.000	3103.000	2827.000
R-hat	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n = 100									
Mean	5.101	1.987	1.404	0.931	-2.068	4.238	1.965	0.944	1.041
Sd	0.302	0.094	0.270	0.369	0.786	0.326	0.100	0.171	0.379
3 HDI	4.528	1.810	0.868	0.330	-3.577	3.651	1.774	0.619	0.365
97 HDI	5.651	2.161	1.878	1.611	-0.572	4.838	2.145	1.242	1.688
ESS(bulk)	1700.000	15623.000	909.000	520.000	7892.000	383.000	13657.000	406.000	321.000
ESS(tail)	5662.000	17727.000	1299.000	1156.000	6741.000	1699.000	16807.000	1239.000	1549.000
R-hat	1.000	1.000	1.000	1.000	1.000	1.010	1.000	1.010	1.020
n = 200									
Mean	4.832	2.025	1.460	0.866	-1.969	4.170	2.023	0.849	1.308
Sd	0.235	0.076	0.238	0.331	0.620	0.226	0.077	0.126	0.264
3 HDI	4.384	1.881	0.968	0.334	-3.152	3.719	1.878	0.629	0.766
97 HDI	5.266	2.169	1.866	1.477	-0.739	4.575	2.165	1.100	1.768
ESS(bulk)	1369.000	15431.000	556.000	335.000	1955.000	364.000	9008.000	335.000	314.000
ESS(tail)	3618.000	16487.000	880.000	943.000	1693.000	443.000	14677.000	519.000	360.000
R-hat	1.000	1.000	1.010	1.020	1.000	1.010	1.000	1.020	1.020
n = 500									
Mean	4.811	1.922	1.311	0.801	-1.418	4.160	1.917	0.918	1.021
Sd	0.221	0.049	0.181	0.266	0.480	0.210	0.050	0.098	0.258
3 HDI	4.383	1.831	0.945	0.344	-2.310	3.751	1.824	0.749	0.508
97 HDI	5.201	2.016	1.614	1.280	-0.449	4.521	2.012	1.101	1.448
ESS(bulk)	1024.000	34358.000	328.000	178.000	2678.000	250.000	13881.000	262.000	232.000
ESS(tail)	4947.000	18812.000	836.000	587.000	4830.000	349.000	16420.000	558.000	315.000
R-hat	1.010	1.000	1.020	1.040	1.000	1.010	1.000	1.010	1.010

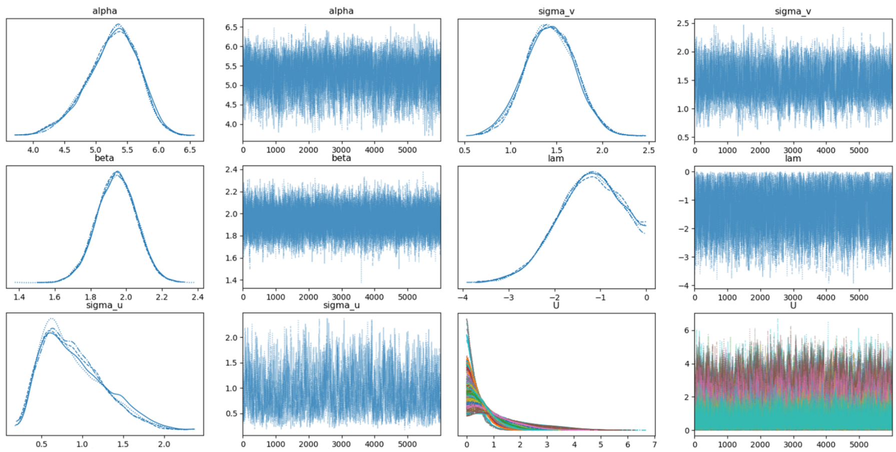
Table 5: Posterior summaries under Skewness (left) and Normal (right) models for $\lambda = -5$

	Skewness Model					Normal Model			
Sample	α	β	σ_v	σ_u	λ	α	β	σ_v	σ_u
True Value	5.000	2.000	1.500	1.000	-5.000	5.000	2.000	1.500	1.000
n = 50									
Mean	4.826	2.143	1.764	0.933	-4.758	3.510	2.205	0.999	1.072
Sd	0.274	0.126	0.321	0.403	0.957	0.432	0.145	0.229	0.452
3 HDI	4.291	1.911	1.109	0.313	-6.581	2.757	1.927	0.558	0.317
97 HDI	5.320	2.385	2.343	1.706	-2.994	4.327	2.474	1.400	1.850
ESS(bulk)	2400.000	11850.000	1533.000	967.000	13503.000	872.000	12485.000	894.000	735.000
ESS(tail)	6905.000	15794.000	1885.000	2113.000	11012.000	2810.000	15597.000	1945.000	2731.000
R-hat	1.000	1.000	1.000	1.000	1.000	1.010	1.000	1.000	1.010
n = 100									
Mean	5.148	1.992	1.433	1.032	-5.023	4.074	1.993	0.840	1.099
Sd	0.169	0.075	0.297	0.367	0.976	0.325	0.092	0.178	0.370
3 HDI	4.818	1.851	0.838	0.394	-6.827	3.473	1.817	0.505	0.396
97 HDI	5.459	2.134	1.933	1.697	-3.139	4.652	2.164	1.150	1.714
ESS(bulk)	2125.000	12433.000	603.000	568.000	16201.000	375.000	13471.000	372.000	344.000
ESS(tail)	5552.000	13651.000	732.000	1014.000	11179.000	1569.000	15568.000	1131.000	1120.000
R-hat	1.000	1.000	1.000	1.000	1.000	1.010	1.000	1.010	1.010
n = 200									
Mean	4.952	2.015	1.507	0.926	-4.659	4.159	2.008	0.676	1.393
Sd	0.159	0.063	0.214	0.305	0.911	0.193	0.069	0.130	0.228
3 HDI	4.644	1.893	1.077	0.396	-6.381	3.793	1.875	0.446	0.957
97 HDI	5.243	2.133	1.875	1.492	-2.975	4.505	2.133	0.926	1.792
ESS(bulk)	947.000	13766.000	556.000	446.000	12898.000	305.000	6670.000	261.000	280.000
ESS(tail)	3369.000	15995.000	839.000	906.000	12382.000	415.000	12948.000	377.000	334.000
R-hat	1.000	1.000	1.010	1.010	1.000	1.010	1.000	1.010	1.010
n = 500									
Mean	5.038	1.944	1.332	0.996	-4.830	4.335	1.949	0.562	1.425
Sd	0.113	0.040	0.189	0.263	0.902	0.094	0.042	0.068	0.108
3 HDI	4.824	1.868	0.974	0.521	-6.556	4.160	1.869	0.439	1.224
97 HDI	5.244	2.019	1.656	1.480	-3.185	4.512	2.026	0.694	1.633
ESS(bulk)	271.000	8612.000	165.000	142.000	3533.000	533.000	7836.000	397.000	485.000
ESS(tail)	350.000	18117.000	383.000	267.000	9316.000	1081.000	15093.000	706.000	926.000
R-hat	1.010	1.000	1.020	1.020	1.000	1.010	1.000	1.010	1.010

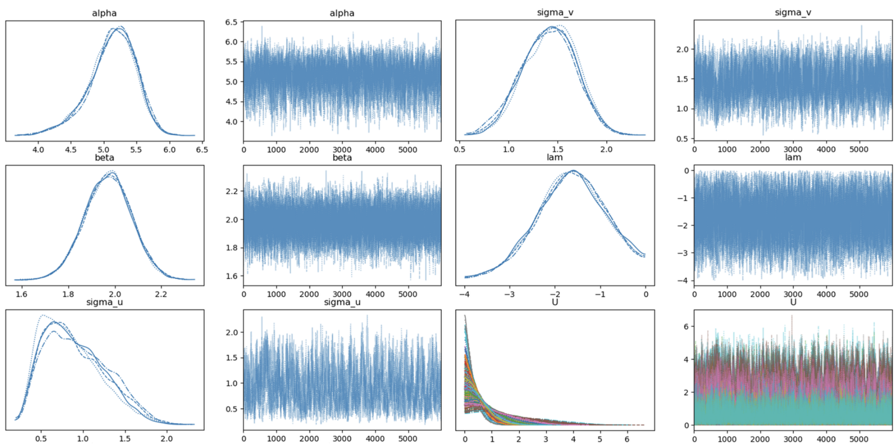
4.2 Traceplots(n=100)



(a) $\lambda = -0.5$

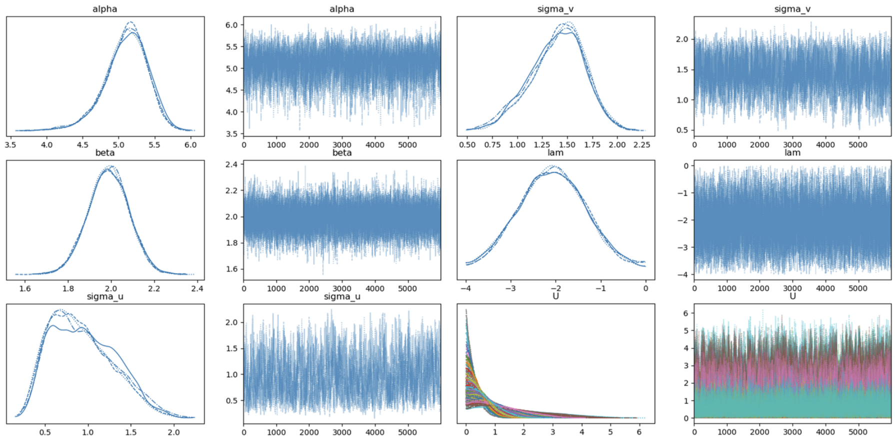


(b) $\lambda = -1$

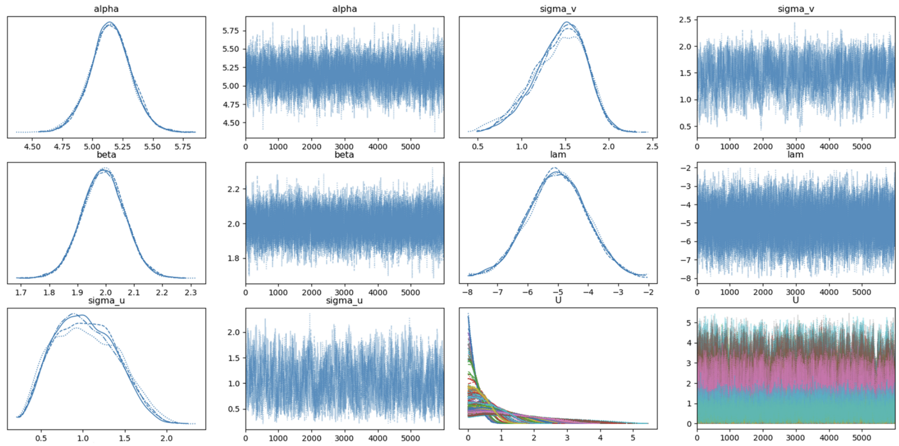


(c) $\lambda = -1.5$

Figure 1: Traceplots of posterior samples for $\lambda = -0.5$, -1, and -1.5 under the skew-normal model.



(d) $\lambda = -2$



(e) $\lambda = -5$

Figure 2: Traceplots of posterior samples for $\lambda = -2$ and -5 under the skew-normal model.

4.3 Model Comparison via LOO: Skew Normal vs Normal (n=100)

Table 6: Leave-one-out cross-validation (LOO) comparison for $\lambda = -0.5$

Model	rank	elpd_loo	p_loo	elpd_diff	weight	se	dse	warning
Normal	0	-174.45	21.55	0.000	1.000	6.53	0.000	FALSE
Skew Normal	1	-174.54	21.00	0.091	0.000	6.58	0.32	FALSE

Table 7: LOO comparison for $\lambda = -1$

Model	rank	elpd_loo	p_loo	elpd_diff	weight	se	dse	warning
Normal	0	-166.31	24.68	0.000	1.000	6.63	0.00	TRUE
Skew Normal	1	-166.61	22.95	0.306	≈ 0.0	6.75	0.53	TRUE

Table 8: LOO comparison for $\lambda = -1.5$

Model	rank	elpd_loo	p_loo	elpd_diff	weight	se	dse	warning
Normal	0	-159.83	30.22	0.000	0.83	6.74	0.00	TRUE
Skew Normal	1	-159.98	25.38	0.147	0.17	6.76	0.67	TRUE

Table 9: LOO comparison for $\lambda = -2$

Model	rank	elpd_loo	p_loo	elpd_diff	weight	se	dse	warning
Skew Normal	0	-155.09	28.83	0.000	1.000	6.65	0.00	TRUE
Normal	1	-155.59	34.44	0.508	0.000	6.65	0.80	TRUE

Table 10: LOO comparison for $\lambda = -5$

Model	rank	elpd_loo	p_loo	elpd_diff	weight	se	dse	warning
Skew Normal	0	-142.70	36.18	0.000	1.000	6.03	0.00	TRUE
Normal	1	-148.33	41.54	5.627	≈ 0.0	6.39	1.84	TRUE

4.4 Posterior Estimates RMSE

Table 11: Posterior estimates RMSE for $\lambda = -5$ across sample sizes and models

Sample Size (n)	**Parameter**	**Skew Model RMSE**	**Normal Model RMSE**
50	α	0.177	1.469
	β	0.144	0.205
	σ_v	0.270	0.510
	σ_u	0.076	0.097
	λ	4.281	–
100	α	0.152	0.928
	β	0.008	0.005
	σ_v	0.072	0.662
	σ_u	0.041	0.100
	λ	4.525	–
200	α	0.054	0.863
	β	0.016	0.007
	σ_v	0.008	0.811
	σ_u	0.080	0.366
	λ	4.140	–
500	α	0.053	0.664
	β	0.057	0.051
	σ_v	0.188	0.939
	σ_u	0.032	0.426
	λ	4.382	–

4.5 Prediction RMSE

Table 12: Root Mean Squared Error (RMSE) for y prediction under $\lambda = -5$ across sample sizes.

Sample Size (n)	**Skew Model RMSE**	**Normal Model RMSE**
50	0.9210	0.7997
100	0.7086	0.6501
200	0.7886	0.4439
500	0.6653	0.3349

4.6 Inefficiency Term RMSE

Table 13: Summary of RMSE for inefficiency term (u) across sample sizes and λ values.

Sample Size (n)	$\lambda = -0.5$		$\lambda = -1$		$\lambda = -1.5$		$\lambda = -2$		$\lambda = -5$	
	Skew	Normal	Skew	Normal	Skew	Normal	Skew	Normal	Skew	Normal
50	0.7671	0.7657	0.7656	0.7698	0.7719	0.7716	0.7711	0.7747	0.7881	0.8202
100	0.7284	0.7294	0.7361	0.7397	0.7406	0.7556	0.7439	0.7715	0.7599	0.7940
200	0.6830	0.7305	0.6917	0.7801	0.6952	0.8284	0.6944	0.8612	0.7037	0.9464
500	0.6606	0.6590	0.6646	0.6616	0.6649	0.6619	0.6618	0.7007	0.6917	0.9834

5 Conclusion

The model proposed by Wei et al. (2025), which utilizes a Bayesian linear framework with a skew-normal half-normal likelihood, shows comparable performance to a standard normal likelihood model across common evaluation metrics such as LOO, RMSE of parameter estimates, and RMSE of observed output predictions. However, for values of the skewness parameter λ that deviate substantially from zero (e.g., -1.5, -2 and -5), the Wei et al. (2025) model exhibits significantly improved accuracy in predicting inefficiency and parameter estimation. This is evidenced by notably lower RMSE values for the posterior mean inefficiency and parameter estimates relative to those produced by the normal model.

References

Aigner, D., Lovell, C., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1):21–37.

Johnes, J. and Portela, M. C. A. S. (2019). The efficiency of higher education institutions: A systematic review of stochastic frontier models. Technical report, Econstor Working Papers.

Salvatier, J., Wiecki, T. V., and Fonnesbeck, C. (2016). Probabilistic programming in python using pymc3. *PeerJ Computer Science*, 2:e55.

Wei, Z., Choy, S. B., Wang, T., and Zhu, X. (2025). Bayesian stochastic frontier models under the skew-normal half-normal settings. *Journal of Productivity Analysis*, pages 1–11.

Worthington, A. C. (2022). A stochastic frontier analysis of queensland public hospital efficiency. Technical Report WP052022, University of Queensland.