Estimating Price Model

Jeong Hyun Lee

December 2022

In this study, we estimate a limited observables discrete choice model in the format of a price model.

Contents

1 Data Generation

1.1 Conditional Multinomial Logit Model

We employ a standard conditional multinomial logit model, with two coefficients representing individual and product characteristics and an intercept.

We set the true coefficients are as follows: α - intercept, γ - price, β - individual.

$$\alpha = 1, \quad \gamma = 0.5 \quad \beta = 0.25$$

For each treatment condition, we take our model with $N = 10000$ individuals and $J = 5$ products (products are indexed from 0, hence $j = 0, ..., 4$)

Following, the data is generated with parameters σ_v^2 and ω., where i is the index for each individual and j is the index for each product.

$$\begin{aligned} X_{ij} &= j + v \quad (v \sim N(0, v)) \\ P_{ij} &= X_{ij} + w \quad (w \sim Uni(-\omega, \omega)) \end{aligned}$$

Then, the utility for each alternative is calculated as

$$U_{ij} = \alpha - \gamma P_{ij} + \beta X_{ij} + \varepsilon_{ij}$$

where $Y_{ij} = 1$ if and only if $j = \arg\max_k U_{ik}$. Alternatively, we can express this as a probability that an individual i with attributes X_{ij} and P_{ij} chooses product j.

$$\pi_{ij}(j \mid X_{ij}, P_{ij}) = P(U_{ij} > U_{ik} \, \forall \, k \neq j \mid \{X_{ij}, P_{ij}\}_{j=0,...,4})$$

Now under the assumption that ε_{ij} have a standard Type I extreme value distribution (which we acknowledge is not the case but for the purpose of the following result),

$$f(\varepsilon) = \exp(-\varepsilon - \exp(-\varepsilon))$$

the probability takes a logit form.

$$\implies P(Y_{ij} = 1 \mid X_{ik}, P_{ik}, \alpha, \gamma, \beta, \forall\{j, k\} \in \{0, \ldots, 4\}) = \frac{\exp(\alpha - \gamma P_{ij} + \beta X_{ij})}{\sum_{k=0}^{4} \exp(\alpha - \gamma P_{ik} + \beta X_{ik})}$$

1.2 Data Specifications

We generate a total of 8 datasets, four with T = 1000 and four with T = 10000 trials. For both sets of trials, we have four cases.

- Base case: we have the variance of the v term as the usual $\sigma_v = 1$ and the range of the uniform distribution for ω as 2.
- High v variance case: we set $\sigma_v = 4$ and ω's specification is the same.
- High ω case: we set the range to 8 and keep $\sigma_v = 1$.
- High v variance and ω: we set $\sigma_v = 4$ and the ω to 8.

2 Treatment Conditions

The purpose of the study is to see if the generated data accurately represents the model given limitations on the observable features. Hence, we aim to estimate the coefficients given the generated data. Thus, we impose three different limitations on the observable features, with one control.

1. Data is the true data (control).

2. Data now omits the real price from non-chosen products for each individuals. Then uses the average price within chosen product j's to replace price for ALL choices, including observed ones.

3. Data omits the real price from non-chosen products for each individuals. Then use the average price within chosen product j' to replace price for *nonchosen* products, and thus observed price is still used for chosen products.

4. Data replaces price for ALL choices with the sum of the predicted price of version 2 and the difference between individual characteristics and product index, i.e. $\hat{P} = \hat{P}_2 + X_{ij} - j$.

3 Summary Statistics

3.1 Trial 1

Table 1: Trial 1 - Base Case

(a) Version 1

	α	γ	β
count	1000	1000	1000
mean	0.013300	0.561330	0.279962
std	0.028868	0.011171	0.013329
min	-0.076258	0.520596	0.236267
25%	-0.006276	0.553548	0.270989
50%	0.014000	0.561429	0.279620
75%	0.032442	0.569114	0.289160
max	0.104443	0.597222	0.320346

(b) Version 2

	α	γ	β
count	1000	1000	1000
mean	0.028462	0.009808	-0.255555
std	0.032360	0.017036	0.011274
min	-0.068681	-0.047651	-0.291430
25%	0.006141	-0.001469	-0.263502
50%	0.027972	0.010240	-0.255180
75%	0.049235	0.021105	-0.247693
max	0.141145	0.061580	-0.220748

(c) Version 3

	α	γ	β
count	1000	1000	1000
mean	-0.038068	-0.073096	-0.303569
std	0.027859	0.011491	0.011491
min	-0.122536	-0.113875	-0.339431
25%	-0.057130	-0.080774	-0.311310
50%	-0.038231	-0.072907	-0.303369
75%	-0.019486	-0.065220	-0.295547
max	0.047718	-0.037045	-0.269500

(d) Version 4

	α	γ	β
count	1000	1000	1000
mean	0.027842	-0.115804	-0.371921
std	0.033341	0.215268	0.208052
min	-0.089238	-1.052591	-1.282735
25%	0.004442	-0.247692	-0.502490
50%	0.027602	-0.114631	-0.372428
75%	0.049612	0.027243	-0.235687
max	0.172250	0.506278	0.224309

Table 2: Trial 1 - High v Variance Case

(a) Version 1

	α	γ	β
count	1000	1000	1000
mean	0.003288	0.553174	0.276633
std	0.026061	0.012084	0.011810
min	-0.079822	0.517175	0.238138
25%	-0.013058	0.545042	0.268678
50%	0.003109	0.553356	0.277169
75%	0.019891	0.561754	0.284645
max	0.091434	0.596400	0.315587

(b) Version 2

	α	γ	β
count	1000	1000	1000
mean	0.011356	0.004542	-0.253136
std	0.035335	0.015648	0.003662
min	-0.106429	-0.047569	-0.266557
25%	-0.013348	-0.006138	-0.255552
50%	0.012021	0.004858	-0.253243
75%	0.036361	0.015500	-0.250583
max	0.133931	0.057501	-0.239500

(c) Version 3

	α	γ	β
count	1000	1000	1000
mean	-0.298754	-0.276753	-0.349026
std	0.030355	0.006983	0.006025
min	-0.396046	-0.298311	-0.369112
25%	-0.319400	-0.281754	-0.353069
50%	-0.297710	-0.276586	-0.349112
75%	-0.279188	-0.272146	-0.344901
max	-0.188499	-0.255408	-0.327618

(d) Version 4

	α	γ	β
count	1000	1000	1000
mean	0.012689	-0.014576	-0.267664
std	0.036314	0.044599	0.043939
min	-0.101792	-0.192345	-0.441745
25%	-0.012077	-0.043750	-0.296179
50%	0.013956	-0.016282	-0.269363
75%	0.037115	0.014971	-0.240088
max	0.130525	0.179188	-0.077445

Table 3: Trial 1 - High ω Case

(a) Version 1

	α	γ	β
count	1000	1000	1000
mean	0.000577	0.542696	0.276518
std	0.041900	0.006590	0.011534
min	-0.124395	0.519544	0.235896
25%	-0.027027	0.537934	0.268984
50%	-0.000046	0.542500	0.276791
75%	0.029262	0.547227	0.283921
max	0.127648	0.565110	0.309507

(b) Version 2

	α	γ	β
count	1000	1000	1000
mean	0.013351	0.004563	-0.138704
std	0.032964	0.018057	0.011318
min	-0.084877	-0.058708	-0.185640
25%	-0.008504	-0.006897	-0.146776
50%	0.013164	0.004388	-0.138574
75%	0.035570	0.016805	-0.131150
max	0.138267	0.059716	-0.107339

(c) Version 3

	α	γ	β
count	1000	1000	1000
mean	0.001091	-0.011574	-0.146916
std	0.028562	0.003595	0.008219
min	-0.076013	-0.023064	-0.180811
25%	-0.018252	-0.013998	-0.152363
50%	0.000447	-0.011607	-0.146946
75%	0.019375	-0.009099	-0.141337
max	0.093379	-0.001404	-0.120960

(d) Version 4

	α	γ	β
count	1000	1000	1000
mean	0.012910	-0.019249	-0.158338
std	0.035215	0.093541	0.086020
min	-0.119783	-0.428180	-0.539010
25%	-0.008297	-0.083527	-0.218893
50%	0.013936	-0.019863	-0.157400
75%	0.035037	0.040119	-0.104886
max	0.147316	0.373712	0.219508

Table 4: Trial 1 - High$\sigma - v$ and ω Case

(a) Version 1

	α	γ	β
count	1000	1000	1000
mean	0.001785	0.540632	0.272837
std	0.035673	0.006665	0.006017
min	-0.120351	0.517335	0.250861
25%	-0.022164	0.536171	0.268835
50%	0.001260	0.540576	0.272734
75%	0.026862	0.545243	0.276913
max	0.112469	0.561456	0.291531

(b) Version 2

	α	γ	β
count	1000	1000	1000
mean	0.011290	0.003476	-0.136925
std	0.035852	0.014886	0.002972
min	-0.109411	-0.042675	-0.146267
25%	-0.013797	-0.006988	-0.138894
50%	0.011195	0.002526	-0.136900
75%	0.035675	0.013119	-0.134899
max	0.115464	0.056587	-0.128427

(c) Version 3

	α	γ	β
count	1000	1000	1000
mean	0.011290	0.003476	-0.136925
std	0.035852	0.014886	0.002972
min	-0.109411	-0.042675	-0.146267
25%	-0.013797	-0.006988	-0.138894
50%	0.011195	0.002526	-0.136900
75%	0.035675	0.013119	-0.134899
max	0.115464	0.056587	-0.128427

(d) Version 4

	α	γ	β
count	1000	1000	1000
mean	0.011412	-0.011165	-0.148065
std	0.037428	0.046773	0.046233
min	-0.117223	-0.206531	-0.342980
25%	-0.013731	-0.041426	-0.178054
50%	0.012108	-0.012296	-0.148865
75%	0.035843	0.018728	-0.118728
max	0.148171	0.148230	0.008074

3.2 Trial 2

Table 5: Trial 2 - Base Case

(a) Version 1

	α	γ	β
count	10000	10000	10000
mean	0.015521	0.561189	0.279116
std	0.029602	0.010958	0.013338
min	-0.113617	0.520596	0.222465
25%	-0.004736	0.553739	0.270253
50%	0.015596	0.561142	0.278978
75%	0.035616	0.568547	0.288077
max	0.132798	0.600366	0.326503

(b) Version 2

	α	γ	β
count	10000	10000	10000
mean	0.029821	0.009423	-0.256322
std	0.032711	0.017094	0.011443
min	-0.102005	-0.049902	-0.301685
25%	0.007603	-0.002056	-0.263969
50%	0.029919	0.009377	-0.256206
75%	0.051648	0.021006	-0.248744
max	0.166142	0.067384	-0.213631

(c) Version 3

	α	γ	β
count	10000	10000	10000
mean	-0.036698	-0.073529	-0.304386
std	0.028669	0.011517	0.011661
min	-0.169096	-0.114190	-0.353458
25%	-0.056062	-0.081389	-0.312141
50%	-0.036679	-0.073530	-0.304344
75%	-0.017234	-0.065672	-0.296633
max	0.087866	-0.032632	-0.260148

(d) Version 4

	α	γ	β
count	10000	10000	10000
mean	0.029155	-0.107671	-0.364594
std	0.034866	0.218374	0.211134
min	-0.107671	-1.243248	-1.470591
25%	0.005792	-0.244169	-0.495315
50%	0.029173	-0.104643	-0.362238
75%	0.051480	0.031607	-0.230554
max	0.235230	0.878673	0.603818

Table 6: Trial 2 - High v Variance Case

(a) Version 1

	α	γ	β
count	10000	10000	10000
mean	0.004219	0.553781	0.277342
std	0.026699	0.012098	0.011884
min	-0.082781	0.509720	0.232830
25%	-0.013778	0.545645	0.269539
50%	0.003953	0.553686	0.277281
75%	0.022344	0.561911	0.285337
max	0.112799	0.602579	0.323809

(b) Version 2

	α	γ	β
count	10000	10000	10000
mean	0.011728	0.004144	-0.253060
std	0.036774	0.016768	0.003683
min	-0.133216	-0.051601	-0.267940
25%	-0.013224	-0.007381	-0.255529
50%	0.011518	0.003789	-0.253051
75%	0.036081	0.015307	-0.250546
max	0.153426	0.080110	-0.238945

(c) Version 3

	α	γ	β
count	10000	10000	10000
mean	-0.298710	-0.277029	-0.348967
std	0.030012	0.006987	0.006084
min	-0.417435	-0.302825	-0.375926
25%	-0.318745	-0.281726	-0.352990
50%	-0.298728	-0.277005	-0.348917
75%	-0.278510	-0.272303	-0.344816
max	-0.176506	-0.251120	-0.324310

(d) Version 4

	α	γ	β
count	10000	10000	10000
mean	0.012404	-0.012619	-0.265629
std	0.036619	0.045437	0.044771
min	-0.152344	-0.215712	-0.463652
25%	-0.011856	-0.043156	-0.295511
50%	0.013194	-0.013355	-0.266561
75%	0.037106	0.016944	-0.236214
max	0.144372	0.184216	-0.072472

Table 7: Trial 2 - High w Case

(a) Version 1

	α	γ	β
count	10000	10000	10000
mean	0.000753	0.542634	0.276002
std	0.041993	0.006568	0.011619
min	-0.149522	0.514082	0.224597
25%	-0.027259	0.538232	0.268098
50%	0.000910	0.542651	0.275957
75%	0.029420	0.546963	0.283809
max	0.158108	0.572318	0.321729

(b) Version 2

	α	γ	β
count	10000	10000	10000
mean	0.014008	0.004481	-0.139181
std	0.033136	0.018216	0.011202
min	-0.104264	-0.058708	-0.185640
25%	-0.008043	-0.007692	-0.146810
50%	0.013943	0.004406	-0.139052
75%	0.036026	0.016626	-0.131554
max	0.141546	0.073214	-0.098480

(c) Version 3

	α	γ	β
count	10000	10000	10000
mean	0.001836	-0.011648	-0.147408
std	0.028789	0.003626	0.008219
min	-0.100352	-0.025591	-0.180811
25%	-0.017405	-0.014111	-0.152970
50%	0.001798	-0.011653	-0.147370
75%	0.020902	-0.009215	-0.141788
max	0.098129	0.001989	-0.116961

(d) Version 4

	α	γ	β
count	10000	10000	10000
mean	0.014197	-0.020764	-0.160240
std	0.035671	0.094760	0.087550
min	-0.131730	-0.622826	-0.747345
25%	-0.008738	-0.082168	-0.216872
50%	0.014415	-0.019618	-0.158833
75%	0.036977	0.040274	-0.104184
max	0.239890	0.381700	0.219508

Table 8: Trial 2 - High σ_v and w Case

(a) Version 1

	α	γ	β
count	10000	10000	10000
mean	0.000626	0.540649	0.272993
std	0.035530	0.006739	0.006178
min	-0.147029	0.515964	0.249582
25%	-0.023364	0.536050	0.268830
50%	0.000620	0.540577	0.273029
75%	0.024434	0.545184	0.277131
max	0.136715	0.564400	0.296938

(b) Version 2

	α	γ	β
count	10000	10000	10000
mean	0.009614	0.003127	-0.136877
std	0.034452	0.014729	0.003049
min	-0.110693	-0.055187	-0.148859
25%	-0.013819	-0.006631	-0.138903
50%	0.009368	0.003059	-0.136839
75%	0.032455	0.012914	-0.134876
max	0.153096	0.068582	-0.125287

(c) Version 3

	α	γ	β
count	10000	10000	10000
mean	-0.102483	-0.084652	-0.157371
std	0.025443	0.003154	0.003556
min	-0.200652	-0.095916	-0.172268
25%	-0.120179	-0.086777	-0.159739
50%	-0.102529	-0.084669	-0.157312
75%	-0.085000	-0.082541	-0.155065
max	-0.006804	-0.072626	-0.143706

(d) Version 4

	α	γ	β
count	10000	10000	10000
mean	0.010009	-0.009900	-0.146766
std	0.035414	0.045101	0.044464
min	-0.161769	-0.209352	-0.345814
25%	-0.013496	-0.038252	-0.174871
50%	0.010131	-0.010383	-0.147148
75%	0.033468	0.019036	-0.118087
max	0.155305	0.254738	0.111907

4 Distribution of Coefficients

4.1 Trial 1

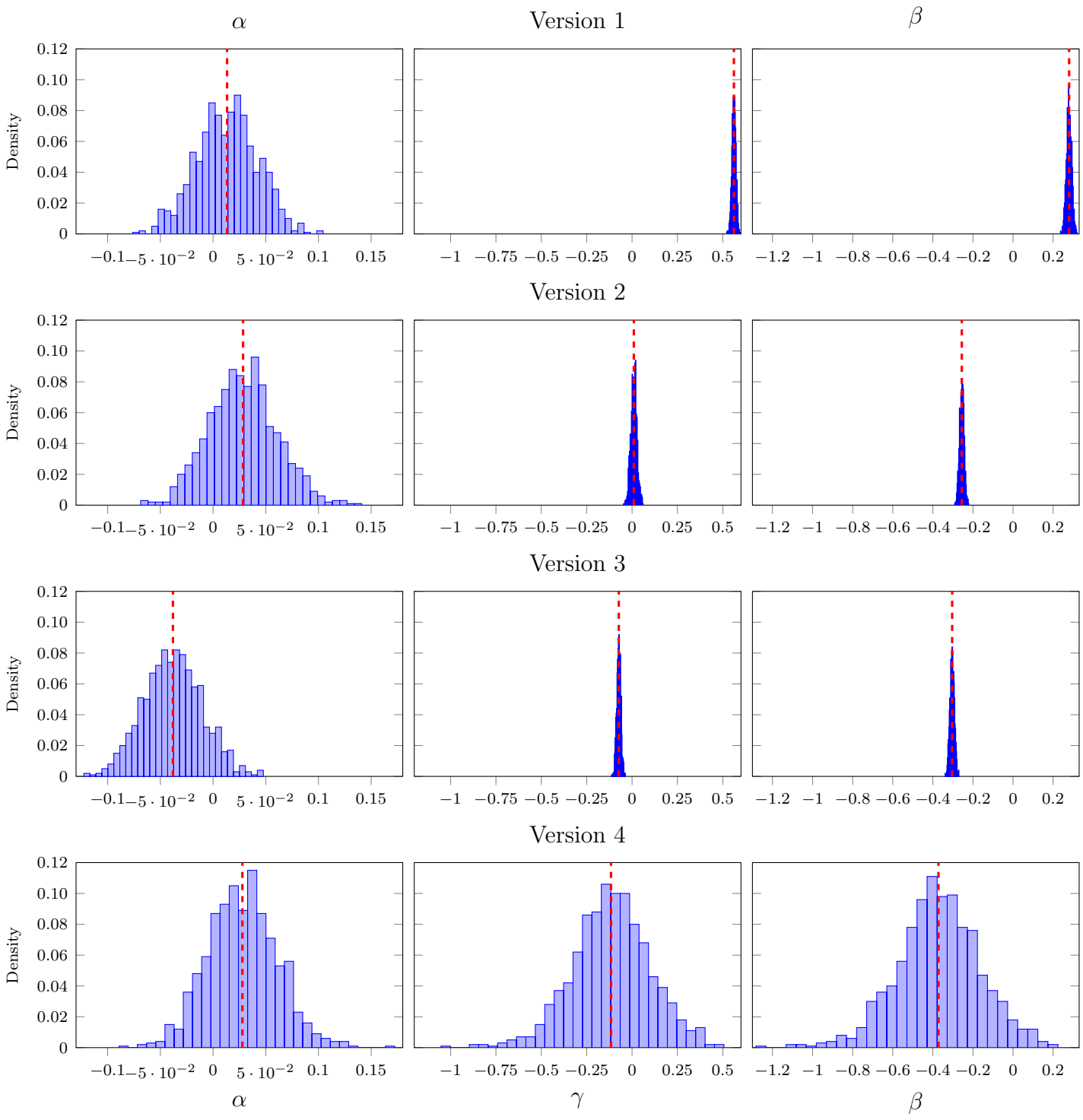

Figure 1: Trial 1 Dist. of Coefficients (Base Case)

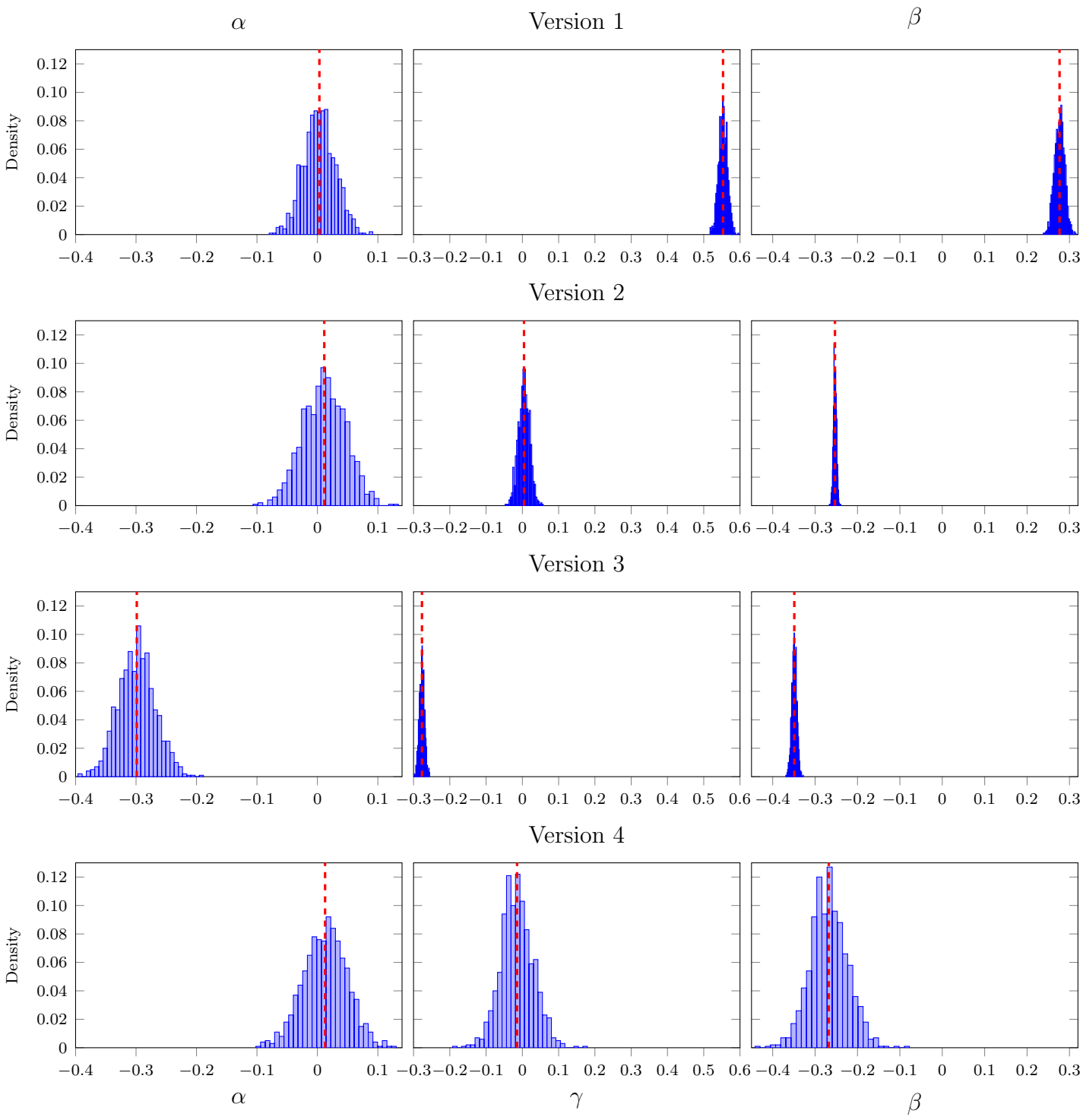

Figure 2: Trial 1 Dist. of Coefficients (High σ_v Case)

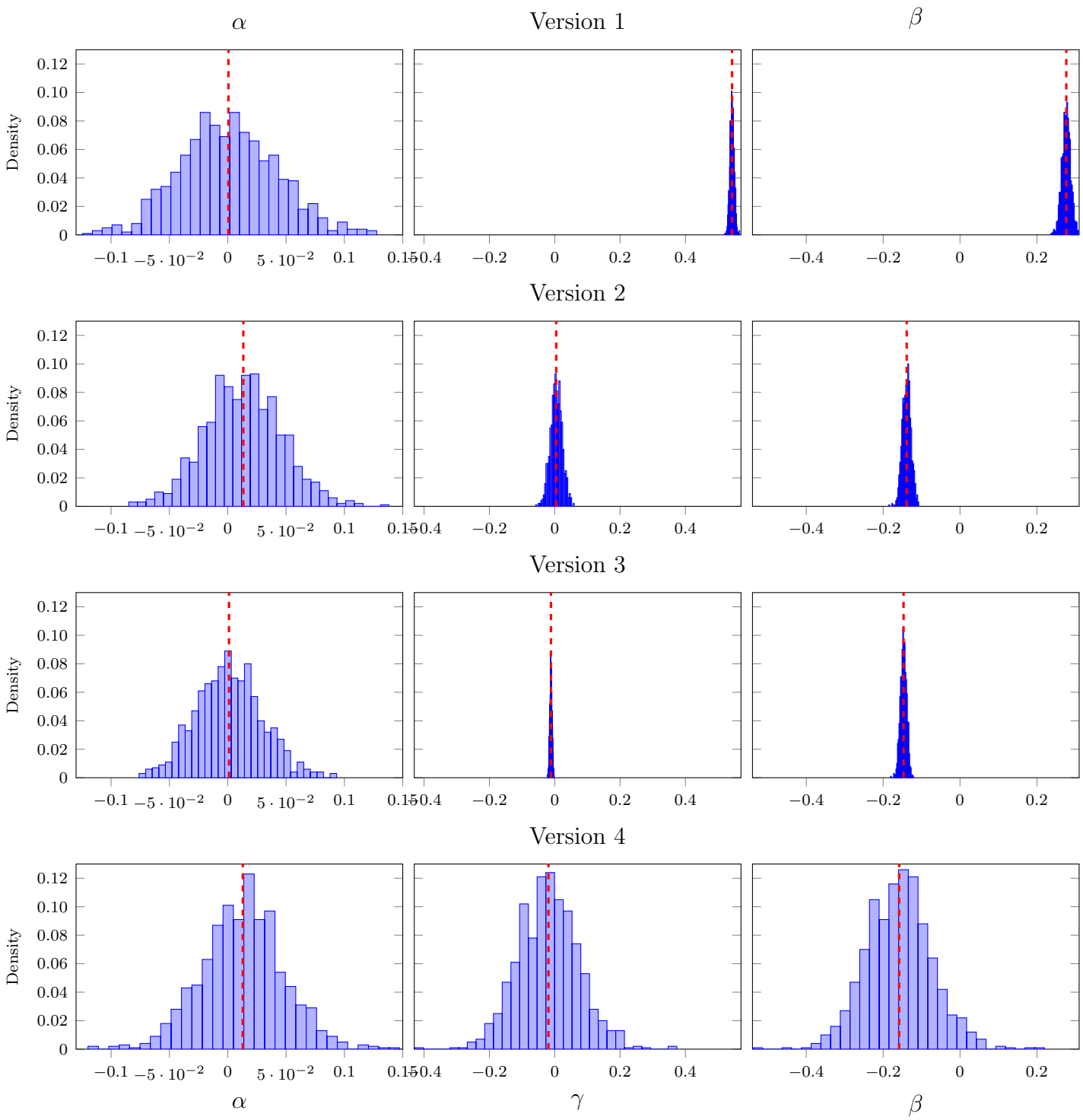

Figure 3: Trial 1 Dist. of Coefficients (High ω Range Case)

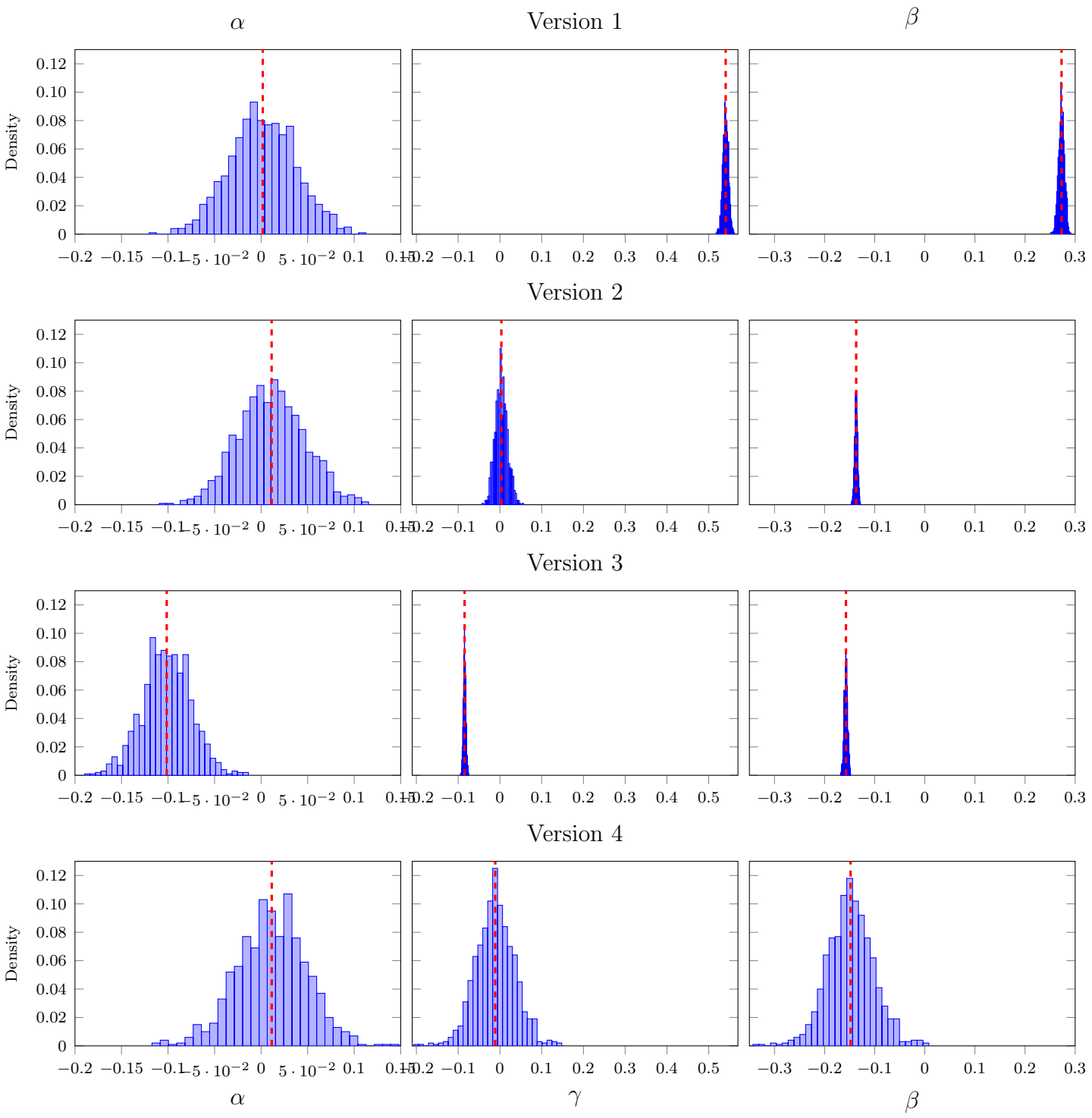

Figure 4: Trial 1 Dist. of Coefficients (High σ_v & ω Range Case)

4.2 Trial 2

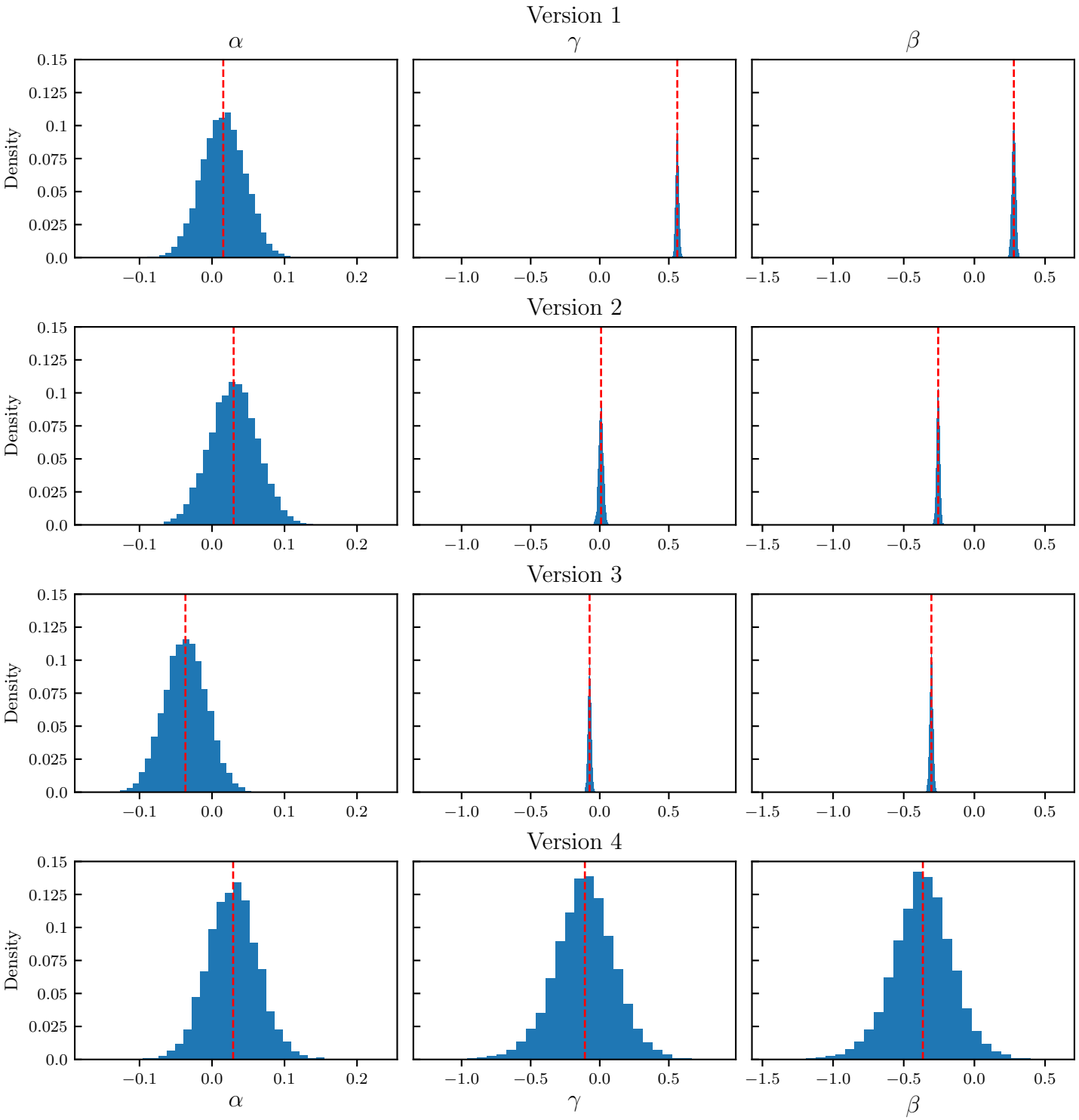

Figure 5: Trial 2 Dist. of Coefficients (Base Case)

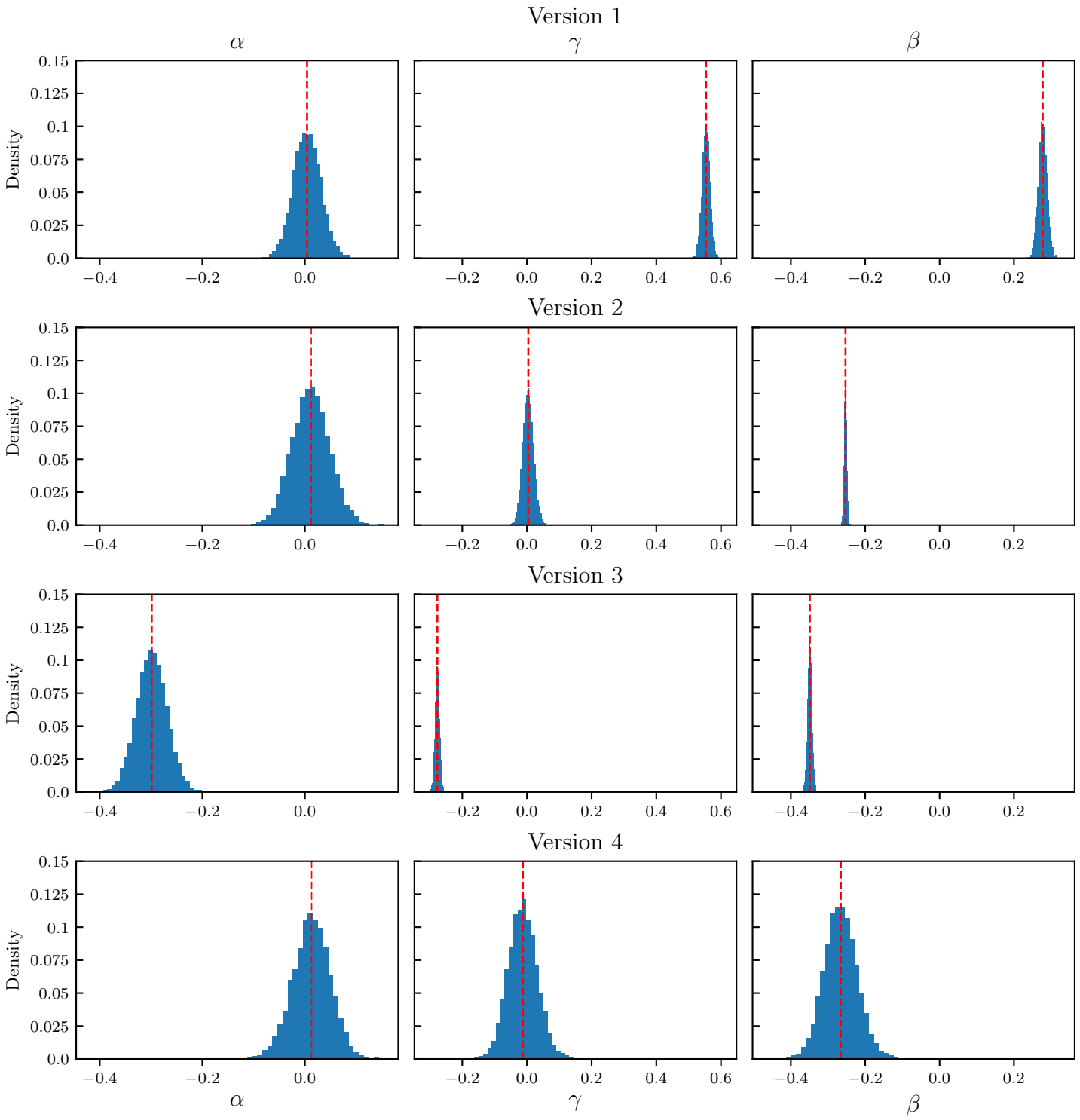

Figure 6: Trial 2 Dist. of Coefficients (High σ_v Case)

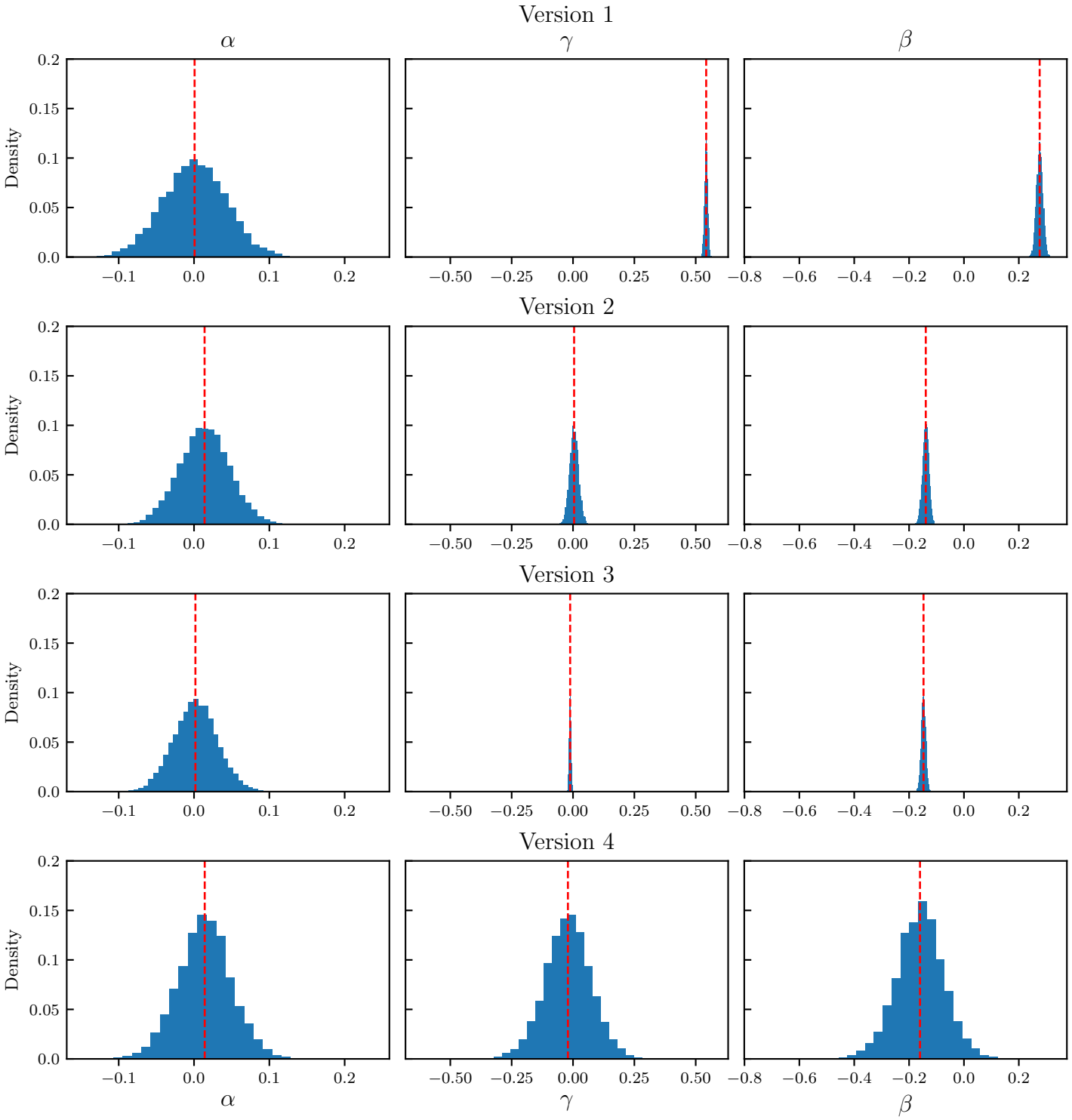

Figure 7: Trial 2 Dist. of Coefficients (High ω Case)

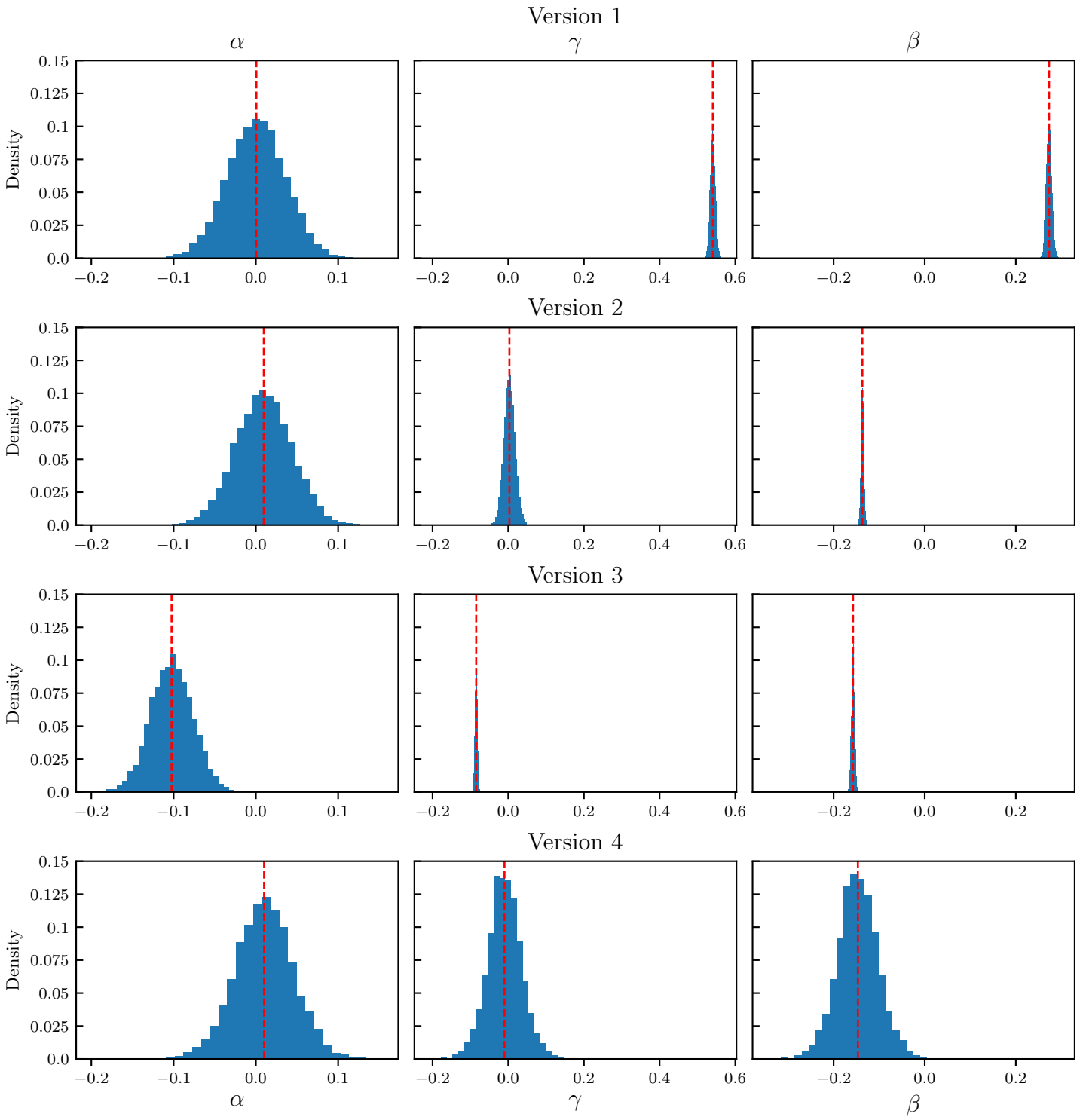

Figure 8: Trial 2 Dist. of Coefficients (High σ_v & ω Case)