A. TECHNICAL PROOFS.

Proof of Theorem 1**.** According to the loss function L_1',

$$E\left[L_1'(\sigma^2_{[1:N]}, \hat{\sigma}^2_{[1:N]})|s^2_{[1:N]}\right] = \sum_i \hat{\sigma}_i^4 E\left[(\sigma_i^2)^{-2}|s^2_{[1:N]}\right] - 2\hat{\sigma}_i^2 E\left[(\sigma_i^2)^{-1}|s^2_{[1:N]}\right] + 1.$$

Consequently,

$$\hat{\sigma}'^2_{i,B} = \frac{E\left[(\sigma_i^2)^{-1}|s^2_{[1:N]}\right]}{E\left[(\sigma_i^2)^{-2}|s^2_{[1:N]}\right]}.$$

For ease of notation, we drop the subscript "i" in the proof. Recall that $p(s^2|\sigma^2)$ is the density function of $s^2|\sigma^2$ and $g(\sigma^2)$ is the prior distribution of σ^2. Note that

$$p(s^2|\sigma^2) = \frac{\left(s^2\right)^{\frac{k}{2}-1} e^{-\frac{ks^2}{2\sigma^2}}}{\Gamma\left(\frac{k}{2}\right) 2^{\frac{k}{2}}} \cdot \left(\frac{k}{\sigma^2}\right)^{\frac{k}{2}}, \quad s^2 > 0.$$

Then

$$f(s^2) = \int p(s^2|\sigma^2) g(\sigma^2) d\sigma^2 = \int C_k \omega(s^2, \sigma^2) d\sigma^2,$$

where

$$C_k = \frac{k^{k/2}}{\Gamma(k/2) 2^{k/2}}, \quad \text{and} \quad \omega(s^2, \sigma^2) = \frac{(s^2)^{k/2-1}}{(\sigma^2)^{k/2}} \exp\left(-\frac{ks^2}{2\sigma^2}\right) g(\sigma^2).$$

Take the derivative of $f(s^2)$ with respect to s^2, we know that

$$\begin{aligned} f'(s^2) &= \int C_k \frac{k-2}{2s^2} \omega(s^2, \sigma^2) d\sigma^2 - \frac{k}{2} \int C_k \frac{1}{\sigma^2} \omega(s^2, \sigma^2) d\sigma^2 \\ &= \frac{k-2}{2s^2} f(s^2) - \frac{k}{2} E\left(\frac{1}{\sigma^2}\middle| s^2\right) \cdot f(s^2). \end{aligned}$$

This leads to

$$\frac{k}{2} E\left(\frac{1}{\sigma^2}\middle| s^2\right) \cdot f(s^2) = \frac{k-2}{2s^2} f(s^2) - f'(s^2). \tag{A1}$$

Take the second order derivative of $f(s^2)$ with respect to s^2, we have

$$\begin{aligned} f''(s^2) &= -\frac{k-2}{2s^4} f(s^2) + \frac{k-2}{2s^2} f'(s^2) - \frac{k}{2} \int C_k \frac{1}{\sigma^2} \left(\frac{k-2}{2s^2} - \frac{k}{2\sigma^2}\right) \omega(s^2, \sigma^2) d\sigma^2 \\ &= -\frac{k-2}{2s^4} f(s^2) + \frac{k-2}{2s^2} f'(s^2) - \frac{k(k-2)}{4s^2} E\left(\frac{1}{\sigma^2}\middle| s^2\right) \cdot f(s^2) + \frac{k^2}{4} E\left(\frac{1}{\sigma^4}\middle| s^2\right) \cdot f(s^2). \end{aligned}$$

Consequently,

$$\frac{k^2}{4} E\left(\frac{1}{\sigma^4}\middle| s^2\right) \cdot f(s^2) = f''(s^2) - \frac{k-2}{s^2} f'(s^2) + \frac{k(k-2)}{4s^4} f(s^2). \tag{A2}$$

Combining (A1) and (A2), we know that

$$\sigma'^2_B = \frac{E\left[(\sigma^2)^{-1}|s^2\right]}{E\left[(\sigma^2)^{-2}|s^2\right]} = \frac{k(k-2)s^2 f(s^2) - 2ks^4 f'(s^2)}{4s^4 f''(s^2) - 4(k-2)s^2 f'(s^2) + k(k-2) f(s^2)}.$$

Proof of Theorem 2**.** For ease of notation, we drop the subscript "i" in the proof. Recall that Stein loss function is defined as

$$L_2(\sigma^2, \hat{\sigma}^2) = \frac{\hat{\sigma}^2}{\sigma^2} - \ln\left(\frac{\hat{\sigma}^2}{\sigma^2}\right) - 1.$$

Consequently,

$$EL_2(\sigma^2, \hat{\sigma}^2|s^2) = \hat{\sigma}^2 E\left[(\sigma^2)^{-1}|s^2\right] - \ln\hat{\sigma}^2 + E(\ln\sigma^2|s^2) - 1.$$

Therefore, the estimator $\hat{\sigma}^2_{Stein}$ which minimizes the above expression is

$$\hat{\sigma}^2_{Stein} = \frac{1}{E\left[(\sigma^2)^{-1}|s^2\right]}$$

According to the proof of Theorem 1,

$$\frac{k}{2}E\left[(\sigma^2)^{-1}|s^2\right] \cdot f(s^2) = \frac{k-2}{2s^2}f(s^2) - f'(s^2).$$

Therefore,

$$\hat{\sigma}^2_{Stein} = \frac{1}{E\left[(\sigma^2)^{-1}|s^2\right]} = \left(\frac{k-2}{ks^2} - \frac{2f'(s^2)}{kf(s^2)}\right)^{-1}.$$

Proof of Theorem 3**.** For ease of notation, we drop the subscript "i" in the proof. Recall that $p(s^2|\sigma^2)$ is the density function of $s^2|\sigma^2$ and $g(\sigma^2)$ is the prior distribution of σ^2. Note that $p(s^2|\sigma^2)$ is given as

$$p(s^2|\sigma^2) = \frac{\left(s^2\right)^{\frac{k}{2}-1} e^{-\frac{ks^2}{2\sigma^2}}}{\Gamma\left(\frac{k}{2}\right)2^{\frac{k}{2}}} \cdot \left(\frac{k}{\sigma^2}\right)^{\frac{k}{2}}, \quad s^2 > 0. \tag{A3}$$

Define $f(s^2)$, $n(s^2)$ and $h(s^2)$ as

$$f(s^2) = \int_0^\infty p(s^2|\sigma^2)g(\sigma^2)\,d\sigma^2, \tag{A4}$$

$$n(s^2) = \int_0^\infty \sigma^2\, p(s^2|\sigma^2)g(\sigma^2)\,d\sigma^2, \tag{A5}$$

and

$$h(s^2) = \int_0^\infty (\sigma^2)^2\, p(s^2|\sigma^2)g(\sigma^2)\,d\sigma^2. \tag{A6}$$

Note that $f(s^2)$ is the marginal distribution of s^2. Then

$$\hat{\sigma}^2_B = \frac{E\left[\,(\sigma^2)^2|s^2\,\right]}{E\left[\,\sigma^2|s^2\,\right]} = \frac{\int_0^\infty(\sigma^2)^2 p(\sigma^2|s^2)\,d\sigma^2}{\int_0^\infty \sigma^2\, p(\sigma^2|s^2)\,d\sigma^2} = \frac{h(s^2)}{n(s^2)}.$$

By differentiating $n(s^2)(s^2)^{-(\frac{k}{2}-1)}$ with respect to s^2, we have

$$\left[n(s^2)(s^2)^{-(\frac{k}{2}-1)}\right]' = -\frac{k}{2}f(s^2)(s^2)^{-(\frac{k}{2}-1)}. \tag{A7}$$

Namely,

$$n(s^2)(s^2)^{-(\frac{k}{2}-1)} = -\frac{k}{2}\int_0^{s^2} f(t)\,t^{-(\frac{k}{2}-1)}dt + C, \tag{A8}$$

for some constant C.

On the other hand, from (A5), the left hand side of (A8) can be expressed as

$$n(s^2)(s^2)^{-(k/2-1)} = \int_0^\infty (s^2)^{-(k/2-1)}\, \sigma^2 p(s^2|\sigma^2) g(\sigma^2)\, d\sigma^2 \tag{A9}$$

$$= \int_0^\infty \frac{(k/2)^{k/2}}{\Gamma\left(k/2\right)} \left(\frac{1}{\sigma^2}\right)^{k/2-1} e^{-\frac{ks^2}{2\sigma^2}}\, g(\sigma^2)\, d\sigma^2.$$

From (A8) and (A9), as s^2 approaches to zero,

$$C = \lim_{s^2\to 0} n(s^2)(s^2)^{-(k/2-1)} = \frac{(k/2)^{k/2}}{\Gamma(k/2)} E\left(\frac{1}{\sigma^2}\right)^{k/2-1} = \frac{k}{2} E\left(\frac{1}{S^2}\right)^{k/2-1},$$

since, for $j = 1, 2$,

$$E\left(\frac{1}{S^2}\right)^{k/2-j} = \frac{(k/2)^{k/2-j}}{\Gamma\left(k/2\right)} E\left(\frac{1}{\sigma^2}\right)^{k/2-j}.$$

Therefore,

$$\begin{aligned} n(s^2)(s^2)^{-(\frac{k}{2}-1)} &= -\frac{k}{2}\int_0^{s^2} f(t)\, t^{-(\frac{k}{2}-1)} dt + \frac{k}{2}\, E\left(\frac{1}{S^2}\right)^{\frac{k}{2}-1} \\ &= -\frac{k}{2}\int_0^{s^2} f(t)\, t^{-(\frac{k}{2}-1)} dt + \frac{k}{2}\int_0^\infty f(t)\, t^{-\left(\frac{k}{2}-1\right)}\, dt \\ &= \frac{k}{2}\int_{s^2}^\infty t^{-(\frac{k}{2}-1)} dF(t). \end{aligned}$$

We can calculate $h(s^2)$ in the similar way. Take the first and second order derivatives of $h(s^2)(s^2)^{-(\frac{k}{2}-1)}$ with respect to s^2, we then have

$$\left[h(s^2)(s^2)^{-(\frac{k}{2}-1)}\right]' = -\frac{k}{2} n(s^2)(s^2)^{-(\frac{k}{2}-1)}, \tag{A10}$$

$$\left[h(s^2)(s^2)^{-(\frac{k}{2}-1)}\right]'' = \frac{k^2}{4} f(s^2)(s^2)^{-(\frac{k}{2}-1)}. \tag{A11}$$

Consequently,

$$\left[h(s^2)(s^2)^{-(\frac{k}{2}-1)}\right]' = \int_0^{s^2} \frac{k^2}{4} f(t) t^{-(\frac{k}{2}-1)}\, dt + C_1, \tag{A12}$$

and

$$h(s^2)(s^2)^{-(\frac{k}{2}-1)} = \int_0^{s^2} \int_0^y \frac{k^2}{4} f(t)\, t^{-(\frac{k}{2}-1)}\, dt\, dy + C_1 s^2 + C_2$$

$$= \frac{k^2}{4}\int_0^{s^2} f(t)\, t^{-(\frac{k}{2}-1)} (s^2 - t)\, dt + C_1 s^2 + C_2, \tag{A13}$$

for some constants C_1 and C_2.

From (A10) and (A12), as s^2 approaches to zero, similar argument shows that

$$C_1 = \lim_{s^2\to 0} \left[h(s^2)(s^2)^{-(\frac{k}{2}-1)}\right]' = -\frac{k^2}{4} E\left(\frac{1}{S^2}\right)^{\frac{k}{2}-1}.$$

Similarly, combine equations (A11) and (A13) and let s^2 approach to zero,

$$C_2 = \lim_{s^2 \to 0} h(s^2)(s^2)^{-(\frac{k}{2}-1)} = \frac{k^2}{4} E\left(\frac{1}{S^2}\right)^{\frac{k}{2}-2}.$$

Thus,

$$\begin{aligned} h(s^2)(s^2)^{-(\frac{k}{2}-1)} &= \frac{k^2}{4}\int_0^{s^2} f(t)\, t^{-(\frac{k}{2}-1)}(s^2-t)\, dt - \frac{k^2}{4} E\left(\frac{1}{S^2}\right)^{\frac{k}{2}-1} s^2 + \frac{k^2}{4} E\left(\frac{1}{S^2}\right)^{\frac{k}{2}-2} \\ &= \frac{k^2}{4}\left[-\left(s^2 \int_{s^2}^{\infty} f(t)\, t^{-\left(\frac{k}{2}-1\right)}\, dt\right) + \left(\int_{s^2}^{\infty} f(t)\, t^{-(\frac{k}{2}-2)}\, dt\right)\right] \\ &= \frac{k^2}{4}\int_{s^2}^{\infty} t^{-(\frac{k}{2}-1)}(t-s^2)\, dF(t). \end{aligned}$$

Therefore,

$$\hat{\sigma}_B^2 = \frac{h(s^2)}{n(s^2)} = \frac{k}{2}\left[\frac{\int_{s^2}^{\infty} t^{-(\frac{k}{2}-2)}\, dF(t)}{\int_{s^2}^{\infty} t^{-(\frac{k}{2}-1)}\, dF(t)} - s^2\right].$$

Proof of Theorem 4**.** We restate one of the momumental theorems in the empirical process, on which our proof is based (Blum, 1955; DeHardt, 1971).

Let $\mathcal{F}$ be a set of measurable function. The bracket $[a, b]$ is the set of all the functions $l \in \mathcal{F}$ with $a \le l \le b$. An ϵ-bracket is a bracket with $\|b-a\| \le \epsilon$. The bracketing number $N_{[]}(\epsilon, \mathcal{F}, L_1(P))$ is the minimum number of ϵ-brackets with which $\mathcal{F}$ can be covered.

Theorem *(Blum-DeHardt) Let $\mathcal{F}$ be a class of measurable functions such that $N_{[]}(\epsilon, \mathcal{F}, L_1(P)) < \infty$, for every $\epsilon > 0$. Then $\mathcal{F}$ is P-Glivenko-Cantelli.*

We only prove the part for the numerator and the denominator can be similarly done. Let $\mathcal{F} = \left\{l_1 : l_1(s^2, u) = (s^2)^{-(k/2-2)}\mathbb{I}(s^2 > u),\ u > 0\right\}$ and $Pl_1(s^2, u) = \int_0^\infty l_1(s^2, u)\, dF(s^2) = \int_u^\infty s^{2-(k/2-2)}\, dF(s^2)$. It suffices to show that $\mathcal{F}$ is a *P-Glivenko-Cantelli* class of functions. Since F is continuous and $\int_0^\infty (s^2)^{-(k/2-2)}\, dF(s^2) < \infty$, for any $\epsilon > 0$, a collection of real numbers $0 = v_0 < v_1 < v_2 < \cdots < v_m = \infty$ can be found such that

$$\begin{aligned} Pl_1(s^2, v_{j-1}) - Pl_1(s^2, v_j) &= \int_{v_{j-1}}^{\infty} (s^2)^{-(k/2-2)}\, dF(s^2) - \int_{v_j}^{\infty} (s^2)^{-(k/2-2)}\, dF(s^2) \\ &= \int_{v_{j-1}}^{v_j} (s^2)^{-(k/2-2)}\, dF(s^2) \\ &\le \epsilon \end{aligned}$$

for all $1 \le j \le m$, with

$$Pl_1(s^2, v_m^-) = \lim_{v_m \uparrow \infty} Pl_1(s^2, v_m) = \lim_{v_m \uparrow \infty} \int_{v_m}^{\infty} (s^2)^{-(k/2-2)}\, dF(s^2) = 0.$$

Consider the collection of brackets $\{[a_j, b_j], 1 \le j \le m\}$, with $a_j(s^2) = s^{2-(k/2-2)}\,\mathbb{I}(s^2 > v_j)$ and $b_j(s^2) = s^{2-(k/2-2)}\,\mathbb{I}(s^2 > v_{j-1})$. Now each $l_1 \in \mathcal{F}$ is in at least one bracket and $|a_j - b_j|_P = Pl_1(s^2, v_{j-1}) - Pl_1(s^2, v_j^-) \le \epsilon$ for all $1 \le j \le m$. Thus, by Blum-DeHardt theorem , $\mathcal{F}$ is a *P-Glivenco-Cantelli Class* of functions.

Proof of Theorem 5**.** Let

$$A_N(s_i^2) = \int_0^\infty l_1(s^2, s_i^2)\, dF_N(s^2), \qquad A(s_i^2) = \int_0^\infty l_1(s^2, s_i^2)\, dF(s^2),$$

and

$$B_N(s_i^2) = \int_0^\infty l_2(s^2, s_i^2)\, dF_N(s^2), \qquad B(s_i^2) = \int_0^\infty l_2(s^2, s_i^2)\, dF(s^2).$$

According to the proof of Theorem 4, $\sup_{s_i^2 \in R} |A_N(s_i^2) - A(s_i^2)| \to 0$ and $\sup_{s_i^2 \in R} |B_N(s_i^2) - B(s_i^2)| \to 0$ $a.s.$. Let $L = \inf_{s_i^2 \in D^\delta} \{B(s_i^2)\}$. Then for any $\epsilon > 0$, when N is sufficiently large

$$\inf_{s_i^2 \in D^\delta} B_N(s_i^2) \geq L - \epsilon \quad a.s.,$$

and $\sup_{s_i^2 \in R} A_N(s_i^2) \leq C,\;\; a.s.$ for some constant C. Then

$$\begin{aligned}
&\sup_{s_i^2 \in D^\delta} \left|\hat{\sigma}^2_{i,F\text{-}EBV} - \hat{\sigma}_{i,B}\right| \\
&= \sup_{s_i^2 \in D^\delta} \left| \frac{A_N(s_i^2)}{B_N(s_i^2)} - \frac{A(s_i^2)}{B(s_i^2)} \right| \\
&= \sup_{s_i^2 \in D^\delta} \left| \frac{A_N(s_i^2)(B(s_i^2) - B_N(s_i^2))}{B_N(s_i^2) B(s_i^2)} + \frac{A_N(s_i^2) - A(s_i^2)}{B(s_i^2)} \right| \\
&\leq \frac{C}{L^2} \sup_{s_i^2 \in D^\delta} \left|B(s_i^2) - B_N(s_i^2)\right| + \frac{1}{L} \sup_{s_i^2 \in D^\delta} \left|A(s_i^2) - A_N(s_i^2)\right| \to 0, \;\; a.s..
\end{aligned}$$

Setting	a	%	s^2	*ELJS*	*TW*	*Smyth*	*mSmyth*	*Vash*	*mVash*	*REBayes*	*Proposed*
		1%	2.53	-0.13	-0.57	-0.67	-0.87	-0.56	-0.87	-0.68	-0.87
I	6	5%	1.95	-0.40	-0.70	-0.66	-0.87	-0.59	-0.87	-0.82	-0.87
		all	0.74	-0.70	-0.71	-0.66	-0.87	-0.66	-0.87	-0.78	-0.86
		1%	2.44	1.01	0.41	-0.06	-0.24	0.81	-0.17	-0.07	-0.22
II	6	5%	1.88	0.57	0.10	-0.08	-0.26	0.64	-0.19	-0.17	-0.26
		all	0.77	-0.09	-0.11	0.02	-0.49	0.12	-0.51	-0.44	-0.54
		1%	2.33	1.02	0.57	-0.14	-0.47	0.72	-0.22	-0.36	-0.49
III	3	5%	1.78	0.57	0.20	-0.14	-0.42	0.57	-0.22	-0.34	-0.44
		all	0.70	-0.05	-0.04	0.10	-0.44	0.17	-0.47	-0.44	-0.61
		1%	2.32	1.06	0.55	-0.22	-0.29	0.69	-0.12	-0.27	-0.34
IV	3	5%	1.77	0.62	0.23	-0.14	-0.28	0.61	-0.13	-0.24	-0.31
		all	0.73	0.02	0.03	0.17	-0.40	0.25	-0.45	-0.38	-0.56

Table 1. *The* $\log_{10}(risk)$ *associated with the loss function (*13*) of the different estimators for the variances under different simulation settings. For each setting, we consider three selection rule: (i) the parameters corresponding to the 1% smallest sample variances; (ii) the parameters corresponding to the 5% smallest sample variances; and (iii) all the parameters.*

B. Additional simulation results

In this section, we include additional simulation results which are not listed in the paper due to the page limit. The numerical results consist of four parts: (a) results of variance estimation post-selection; and (b) results of Finite Bayes inference problem.

(a) Results of variance estimation post-selection.

To help the readers, we restate the simulation settings here. Let σ_i^2's be the parameters, and the sample variances s_i^2's are generated according to Model 1 where the degrees of freedom k is chosen as 5. We consider the following different choices of the prior $g(\sigma^2)$:

Setting I: $\sigma_i^2 \sim$ inverse gamma distribution: $IG(a, 1)$ where $a = 10$ and 6;
Setting II: $\sigma_i^2 \sim$ Mixture of inverse gamma distributions: $0.2IG(a, 1) + 0.4IG(8, 6) + 0.4IG(9, 19)$, where $a = 10$ and 6;
Setting III: $\sigma_i^2 = a$ with 0.4 probability and $1/a$ with 0.6 probability, where $a = 3$ and 4;
Setting IV: $\sigma_i^2 \sim$ Mixture of inverse Gaussian distributions: $0.4InvGauss(1/a, 1) + 0.6InvGauss(a, a^4)$, where $a = 2$ and 3.

After generating the data, order the sample variances increasingly. We consider three different selection rules: (i) select the parameters corresponding to the 1% smallest sample variances; (ii) select the parameters corresponding to the 5% smallest sample variances; and (iii) all the parameters. We report $\log_{10}(risk)$ in Table 1.

(b) Results of finite Bayes inference problem.

Next, we consider the *finite Bayes inference* problem. Namely, for each generated data set $\boldsymbol{s}^2$ and a new observation s_0^2, we calculate the estimated values based on different approaches and calculate the loss according to the loss function (9). We calculate the risk based on 500 replications and reported the results in Table 2.

Setting	(a, b)	s^2	*ELJS*	*TW*	*Smyth*	*mSmyth*	*Vash*	*mVash*	*REBayes*	*Proposed*
I	6	0.3	0.07	-0.86	-0.81	-1	-0.8	-1	-0.91	-0.98
II	6	0.64	0.43	-0.18	-0.04	-0.53	-0.02	-0.54	-0.52	-0.59
III	3	0.92	0.72	-0.02	0.06	-0.46	0.18	-0.48	-0.54	-0.61
IV	3	0.43	0.21	-0.08	0.08	-0.43	0.16	-0.46	-0.44	-0.59

Table 2. *The $\log_{10}(risk)$ associated with the loss function (9) of the different estimators for the finite Bayes inference problem.*