A Appendix to "Measuring Prior Sensitivity and Prior Informativness in Large Bayesian Models"

A.1 Derivation of PI from Axiomatic Requirements

Consider functions f_k that map the triple (v, J, Σ_p) to the unit interval, $f_k(v, J, \Sigma_p) \in [0, 1]$. Under a linear reparamterization $\theta^* = H\theta$, the parameter of interest $v'\theta$ becomes $v'\theta = (H^{-1\prime}v)'\theta^* = v^{*\prime}\theta^*$ with $v^* = H^{-1\prime}v$. Denote the implied prior and posterior of θ^* by $p^*(\theta^*) = |H|^{-1}p(H^{-1}\theta^*)$ and π^*, respectively, so that $E_{p^*}[\theta^*] = \mu_{p^*} = HE_p[\theta]$ and $\Sigma_p^* = H\Sigma_p H'$. Let $\mu_\pi^*(\alpha^*)$ be the posterior mean of θ^* under the prior (10), where $\alpha^* = H\alpha$. By a change of variables and the chain rule, we obtain

$$J^* = \frac{\partial \mu_\pi^*(\alpha^*)}{\partial \alpha^{*\prime}}|_{\alpha^*=0} = HJH^{-1} = H\Sigma_\pi \Sigma_p^{-1} H^{-1}. \tag{36}$$

Thus, invariance to linear reparametrizations formally corresponds to

Condition 1 $f_k(v, J, \Sigma_p) = f_k(v^*, J^*, \Sigma_{p^*}) = f_k(H^{-1\prime}v, H\Sigma_\pi\Sigma_p^{-1}H^{-1}, H\Sigma_p H')$ *for all full rank matrices* H.

As a special case, let $H = DQ'P$, were P' is the Choleksy decomposition of Σ_p^{-1}, the columns of Q are the normalized eigenvectors of $P\Sigma_\pi P'$, and D is the diagonal matrix with diagonal elements equal to -1 if the corresponding element in $Q'P^{-1\prime}v$ is negative, and equal to one otherwise. Then $\Sigma_{p^*} = I_k$, J^* is diagonal with $J^* = \text{diag}(\lambda_1, \cdots, \lambda_k)$, and v^* has nonnegative elements. The problem is thus effectively reduced to identifying a suitable function $g_k : \mathbb{R}^{2k} \mapsto [0, 1]$ that maps

$$\left(\begin{pmatrix} \omega_1^2 \\ \lambda_1 \end{pmatrix}, \cdots, \begin{pmatrix} \omega_k^2 \\ \lambda_k \end{pmatrix} \right) \tag{37}$$

with $v^* = (\omega_1, \cdots, \omega_k)'$ to the unit interval. Note that Condition 1 also implies that g_k is invariant to permutations of the k bivariate vectors $(\omega_i^2, \lambda_i)'$, $i = 1, \cdots, k$, as the order of the eigenvectors in Q can be chosen arbitrarily. The diagonal elements of J^* are recognized as the eigenvalues of the matrix J, since $J^* = HJH^{-1}$ implies that J^* and J are similar.

The second and third set of constraints of the main text now corresponds to the following conditions on g_k.

Condition 2 *For any integers k and $m < k$, and any values of $\{\{\omega_i, \lambda_i\}_{i=1}^k\}$:*

(a) $g_1\left(\begin{pmatrix} \omega_1^2 \\ \lambda_1 \end{pmatrix} \right) = \min(\lambda_1, 1)$*;*

(b) $g_{k+1}\left(\begin{pmatrix} \omega_1^2 \\ \lambda_1 \end{pmatrix}, \cdots, \begin{pmatrix} \omega_k^2 \\ \lambda_k \end{pmatrix}, \begin{pmatrix} 0 \\ \lambda_{k+1} \end{pmatrix} \right) = g_k\left(\begin{pmatrix} \omega_1^2 \\ \lambda_1 \end{pmatrix}, \cdots, \begin{pmatrix} \omega_k^2 \\ \lambda_k \end{pmatrix} \right)$*;*

(c) $g_k\left(\begin{pmatrix} \omega_1^2 \\ \lambda_1 \end{pmatrix}, \cdots, \begin{pmatrix} \omega_k^2 \\ \lambda_k \end{pmatrix} \right)$ *has range* $[0, 1]$*, is weakly increasing in* λ_1*, and, for* $\omega_1^2 > 0$ *and* $\max_{i \le k} \lambda_i < 1$*, is continuous in* (ω_1^2, λ_1) *and strictly increasing and differentiable in* λ_1*;*

(d) $g_k\left(\begin{pmatrix}\omega_1^2\\ \lambda_1\end{pmatrix},\cdots,\begin{pmatrix}\omega_{k-2}^2\\ \lambda_{k-2}\end{pmatrix},\begin{pmatrix}\omega_{k-1}^2\\ \lambda_k\end{pmatrix},\begin{pmatrix}\omega_k^2\\ \lambda_k\end{pmatrix}\right)=g_k\left(\begin{pmatrix}\omega_1^2\\ \lambda_1\end{pmatrix},\cdots,\begin{pmatrix}\omega_{k-2}^2\\ \lambda_{k-2}\end{pmatrix},\begin{pmatrix}\omega_{k-1}^2+\omega_k^2\\ \lambda_k\end{pmatrix},\begin{pmatrix}0\\ \lambda_k\end{pmatrix}\right)$;

(e) $g_k\left(\begin{pmatrix}\omega_1^2\\ \lambda_1\end{pmatrix},\cdots,\begin{pmatrix}\omega_k^2\\ \lambda_k\end{pmatrix}\right) = g_k\left(\begin{pmatrix}\omega_1^2\\ \bar\lambda_m\end{pmatrix},\cdots,\begin{pmatrix}\omega_m^2\\ \bar\lambda_m\end{pmatrix},\begin{pmatrix}\omega_{m+1}^2\\ \lambda_{m+1}\end{pmatrix},\cdots,\begin{pmatrix}\omega_k^2\\ \lambda_k\end{pmatrix}\right)$ *for* $\bar\lambda_m = g_m\left(\begin{pmatrix}\omega_1^2\\ \lambda_1\end{pmatrix},\cdots,\begin{pmatrix}\omega_m^2\\ \lambda_m\end{pmatrix}\right)$.

Condition 3 *For* $\lambda_1 < 1$, $g_2\left(\begin{pmatrix}1\\ \lambda_1\end{pmatrix},\begin{pmatrix}1\\ 0\end{pmatrix}\right)=\dfrac{\lambda_1}{2-\lambda_1}$.

The main theoretical result is formally stated as follows:

Theorem 2 *Under Conditions 1–3,*

$$g_k\left(\begin{pmatrix}\omega_1^2\\ \lambda_1\end{pmatrix},\cdots,\begin{pmatrix}\omega_k^2\\ \lambda_k\end{pmatrix}\right)=\begin{cases}1 & \text{if } (\max_{i\le k}\omega_i^2\mathbf{1}[\lambda_i\ge 1])>0\\ 1-\dfrac{\sum_{i=1}^k\omega_i^2}{\sum_{i=1}^k\frac{\omega_i^2}{1-\lambda_i}}=1-\dfrac{v^{*\prime}v^*}{v^{*\prime}(I_k-J^*)^{-1}v^*} & \text{otherwise}\end{cases}$$

The expression (17) for PI in the main text now follows by substituting $v^*=H^{-1\prime}v$ and $J^*=HJH^{-1}$.

Proof. By Condition 2 (b) and the permutation invariance noted below Condition 1, we can restrict attention to the case where $\omega_i^2>0$ for all $i=1,\cdots,k$. Consider first the case of overall limited prior informativeness in the sense of Definition 1, that is $\lambda_i<1$ for $i=1,\cdots,k$. We start by showing that Conditions 1 and 2 imply the four Axioms of Kitagawa (1934), with g_k, k and (ω_i^2,λ_i) playing the role of M_n, n and (w_i,x_i) in Kitagawa's notation. As noted, Axiom 1 follows from Condition 1. Condition 2 (d) implies Axiom 2. Repeated application of Condition 2 (d) yields $g_k\left(\begin{pmatrix}\omega_1^2\\ \lambda_k\end{pmatrix},\begin{pmatrix}\omega_2^2\\ \lambda_k\end{pmatrix},\cdots,\begin{pmatrix}\omega_k^2\\ \lambda_k\end{pmatrix}\right)=g_1\left(\begin{pmatrix}\sum_{i=1}^k\omega_i^2\\ \lambda_k\end{pmatrix}\right)$, which equals $\min(\lambda_k,1)=\lambda_k$ by Condition 2 (a). This shows that Axiom 3 is satisfied. Finally, for Axiom 4, note that applying the permutation invariance and Condition 2 (e) and (b) (repeatedly) yields $g_k\left(\begin{pmatrix}\omega_1^2\\ \lambda_1\end{pmatrix},\cdots,\begin{pmatrix}\omega_k^2\\ \lambda_k\end{pmatrix}\right)=g_k\left(\begin{pmatrix}\omega_1^2\\ \bar\lambda_m\end{pmatrix},\cdots,\begin{pmatrix}\omega_m^2\\ \bar\lambda_m\end{pmatrix},\begin{pmatrix}\omega_{m+1}^2\\ \lambda_{m+1}\end{pmatrix},\cdots\begin{pmatrix}\omega_k\\ \lambda_k\end{pmatrix}\right)=g_{k-m+1}\left(\begin{pmatrix}\omega_{m+1}^2\\ \lambda_{m+1}\end{pmatrix},\cdots,\begin{pmatrix}\omega_k^2\\ \lambda_k\end{pmatrix},\begin{pmatrix}\sum_{i=1}^m\omega_m^2\\ \bar\lambda_m\end{pmatrix}\right)$, so that Axiom 4 follows from another application of Condition 2 (b), with Kitagawa's w_r^* equal to $w_r^*=\sum_{i=1}^r w_i$.

Thus, Kitagawa's results are applicable and imply that g_k is of the form

$$g_k\left(\begin{pmatrix}\omega_1^2\\ \lambda_1\end{pmatrix},\cdots,\begin{pmatrix}\omega_k^2\\ \lambda_k\end{pmatrix}\right)=\phi^{-1}\left(\frac{\sum_{i=1}^k\omega_i^2\phi(\lambda_i)}{\sum_{i=1}^k\omega_i^2}\right)\tag{38}$$

where $\phi:[0,1)\mapsto\mathbb{R}$ is a strictly monotone increasing, continuous function with strictly monotone increasing and continuous inverse ϕ^{-1} (the continuity is not asserted by Kitagawa, but follows from

Kolmogorov's (1930) Theorem invoked in Kitagawa's proof). Without loss of generality, normalize $\phi(0) = 1$.

We now show that ϕ is differentiable at 0. Recall that every strictly monotone function is Lebesgue almost everywhere differentiable. Thus, the two $[0,1) \mapsto \mathbb{R}$ functions $\chi(\lambda) = \frac{1}{2}\phi(0)+\frac{1}{2}\phi(\lambda)$ and $\phi^{-1}(\chi(\lambda))$ are almost everywhere differentiable. Pick $\lambda_0 > 0$ such that both are differentiable at $\lambda = \lambda_0$. We first argue that this implies that ϕ^{-1} is differentiable at $x_0 = \chi(\lambda_0)$. Let h_n be arbitrary nonzero numbers converging to zero as $n \to \infty$. By continuity and monotonicity of χ, there exists, for all large enough n, $h'_n \neq 0$ such that $h_n = \chi(\lambda_0 + h'_n) - \chi(\lambda_0)$, and $h'_n \to 0$. Thus

$$\begin{aligned}
\Delta &= \lim_{n\to\infty} \frac{\phi^{-1}(x_0 + h_n) - \phi^{-1}(x_0)}{h_n} \qquad (39)\\
&= \lim_{n\to\infty} \frac{\phi^{-1}(\chi(\lambda_0 + h'_n)) - \phi^{-1}(\chi(\lambda_0))}{\chi(\lambda_0 + h'_n) - \chi(\lambda_0)}\\
&= \lim_{n\to\infty} \frac{\phi^{-1}(\chi(\lambda_0 + h'_n)) - \phi^{-1}(\chi(\lambda_0))}{h'_n} \cdot \frac{h'_n}{\chi(\lambda_0 + h'_n) - \chi(\lambda_0)}\\
&= \frac{d\phi^{-1}(\chi(\lambda))}{d\lambda}|_{\lambda=\lambda_0} / \frac{d\chi(\lambda)}{d\lambda}|_{\lambda=\lambda_0}
\end{aligned}$$

by the product rule for limits, so that ϕ^{-1} is differentiable at $x_0 = \chi(\lambda_0)$. Furthermore, by the continuity of ϕ at 0,

$$\frac{\phi^{-1}\left(\frac{1}{2}\phi(h_n) + \frac{1}{2}\phi(\lambda_0)\right) - \phi^{-1}(\frac{1}{2}\phi(0) + \frac{1}{2}\phi(\lambda_0))}{h_n} = \Delta_n \frac{\frac{1}{2}\phi(h_n) - \frac{1}{2}\phi(0)}{h_n} \qquad (40)$$

where $\Delta_n \to \Delta$ as $n \to \infty$. By Condition 2 (c), $g_2\left(\begin{pmatrix}1\\ \lambda_1\end{pmatrix}, \begin{pmatrix}1\\ \lambda_0\end{pmatrix}\right) = \phi^{-1}\left(\frac{1}{2}\phi(\lambda_1) + \frac{1}{2}\phi(\lambda_0)\right)$ is differentiable in λ_1 at $\lambda_1 = 0$. Thus, the limit of (40) as $n \to \infty$ exists and doesn't depend on h_n, which implies differentiability of ϕ at 0.

Now by Condition 3,

$$g_2\left(\begin{pmatrix}1\\ \lambda_1\end{pmatrix}, \begin{pmatrix}1\\ 0\end{pmatrix}\right) = \phi^{-1}\left(\tfrac{1}{2}\phi(\lambda_1) + \tfrac{1}{2}\phi(0)\right) = \frac{\lambda_1}{2-\lambda_1}. \qquad (41)$$

Define the continuous and strictly monotone increasing function $\varphi : [0,1) \mapsto \mathbb{R}$ as $\varphi(\lambda) = 1/(1-\lambda)$, and let $h : [1,\infty) \mapsto \mathbb{R}$ be the monotone increasing function such that $\phi(\lambda) = h(\varphi(\lambda))$. The monotonicity and differentiability of ϕ at zero then implies that $h(x)$ has a positive derivative at $x = 1$. Furthermore, $h(1) = 1$ by the normalization $\phi(0) = 1$, and using (41), we have $\phi(\lambda_1)+\phi(0) = 2\phi(\lambda_1/(2-\lambda_1))$ for all $\lambda_1 \in [0,1)$, so that

$$h\left(\frac{1}{1-\lambda_1}\right) + 1 = 2h\left(\frac{2-\lambda_1}{2-2\lambda_1}\right). \qquad (42)$$

With $\lambda_1 = 1 - 1/x$, we obtain $h(x) + 1 = 2h((x+1)/2)$ for all $x \in [1, \infty)$. Repeated substitution yields $h(x) - 1 = 2^j h(2^{-j}x + 1 - 2^{-j}) - 2^j$ for all integer j, so that for $x_1, x_2 \in (1, \infty)$

$$\frac{h(x_1)-1}{h(x_2)-1} = \frac{x_1 - 1}{x_2 - 1}\frac{\frac{h(1+2^{-j}(x_1-1))-1}{2^{-j}(x_1-1)}}{\frac{h(1+2^{-j}(x_2-1))-1}{2^{-j}(x_2-1)}} = \frac{x_1-1}{x_2-1}\frac{dh(x)/dx|_{x=1}}{dh(x)/dx|_{x=1}} = \frac{x_1-1}{x_2-1}. \tag{43}$$

Thus h is linear function, so that $\phi(\lambda) = \varphi(\lambda) = 1/(1-\lambda)$, and the result follows.

Finally, consider the case where $\lambda_i \geq 1$ for some i. Let $\tilde{\lambda}_i(n) = 1 - h_n$ if $\lambda_i \geq 1$ and $\tilde{\lambda}_i(n) = \lambda_i$ otherwise, where h_n is a positive sequence converging to zero. Applying the result for the overall identified case, we obtain $\lim_{n\to\infty} g_k\left(\begin{pmatrix}\omega_1^2 \\ \tilde{\lambda}_1(n)\end{pmatrix}, \cdots, \begin{pmatrix}\omega_k^2 \\ \tilde{\lambda}_k(n)\end{pmatrix}\right) = 1$. Furthermore, by permutation invariance and Condition 2 (c), $g_k\left(\begin{pmatrix}\omega_1^2 \\ \lambda_1\end{pmatrix}, \cdots, \begin{pmatrix}\omega_k^2 \\ \lambda_k\end{pmatrix}\right) \geq g_k\left(\begin{pmatrix}\omega_1^2 \\ \tilde{\lambda}_1(n)\end{pmatrix}, \cdots, \begin{pmatrix}\omega_k^2 \\ \tilde{\lambda}_k(n)\end{pmatrix}\right)$ for all n, so that the result follows from the range upper bound in Condition 2 (c). ■

A.2 Proof of Inequalities of Section 3.1

Note that with $H = DQ'P$ (as discussed below Condition 1 above), for any vector $v^* = H^{-1\prime}v$, we have $v'\Sigma_p v = v^{*\prime}v^* = \sum_{i=1}^k \omega_i^2$, $v'\Sigma_\pi v = v^{*\prime}J^*v^* = \sum_{i=1}^k \omega_i^2\lambda_i$, $v'\Sigma_\pi\Sigma_p^{-1}\Sigma_\pi v = v^{*\prime}J^{*2}v^* = \sum_{i=1}^k \omega_i^2\lambda_i^2$ and, for $\lambda_{\max} < 1$, $\text{PI} = \varphi^{-1}(\sum_{i=1}^k \omega_i^2\varphi(\lambda_i)/\sum_{i=1}^k \omega_i^2)$ with $\varphi(\lambda) = 1/(1-\lambda)$.

Inequality (18) follows from $\sum_{i=1}^k \omega_i^2\lambda_i^2 \leq \lambda_{\max}\sum_{i=1}^k \omega_i^2\lambda_i$;

(19) follows from $\sum_{i=1}^k \omega_i^2\lambda_i^2/\sum_{i=1}^k \omega_i^2 \geq (\sum_{i=1}^k \omega_i^2\lambda_i/\sum_{i=1}^k \omega_i^2)^2$ by convexity;

(20) follows from $\sum_{i=1}^k \omega_i^2\varphi(\lambda_i)/\sum_{i=1}^k \omega_i^2 \geq \varphi(\sum_{i=1}^k \omega_i^2\lambda_i/\sum_{i=1}^k \omega_i^2)$ by convexity of φ;

(21) follows from $\sum_{i=1}^k \omega_i^2\varphi(\lambda_i) \leq \sum_{i=1}^k \omega_i^2\varphi(\lambda_{\max})$ for $\lambda_{\max} < 1$, and the inequality is trivial otherwise;

for (22), note that $\text{PS}/\sqrt{v'\Sigma_p v} = \varphi_{\text{PS}}^{-1}(\sum_{i=1}^k \omega_i^2\varphi_{\text{PS}}(\lambda_i)/\sum_{i=1}^k \omega_i^2)$ with $\varphi_{\text{PS}}(x) = x^2$. Both PI and $\text{PS}/\sqrt{v'\Sigma_p v}$ can thus be considered the certainty equivalence of an expected utility maximizer with utility function φ and φ_{PS}, respectively, facing a lottery with payoff's $\{\lambda_i\}_{i=1}^k$ with probabilities $\{\omega_i^2/\sum_{j=1}^k \omega_j^2\}_{i=1}^k$. The result now follows from Pratt's (1964) Theorem 1, since a calculation shows that φ has a weakly larger (negative) coefficient of absolute risk aversion than φ_{PS} on the interval $[0, 1/3]$.

Inequality (23) follows from $v^{*\prime}(I_k - J^*)^{-1}v^* = v^{*\prime}\sum_{i=0}^\infty (J^*)^i v^* \geq v^{*\prime}(I + J^* + J^{*2})v^*$, so that $\text{PI} = 1 - v^{*\prime}v^*/v^{*\prime}(I_k - J^*)^{-1}v^* \geq v^{*\prime}(J^* + J^{*2})v^*/v^{*\prime}(I + J^* + J^{*2})v^* \geq \frac{2}{3}v^{*\prime}J^{*2}v^*/v^{*\prime}v^*$.