File S1

A technical note on Vuong's test

Vuong (1989) fully characterized the asymptotic distribution of the log-likelihood ratio statistic under the most general conditions. He showed that the form of the asymptotic distribution of the log-likelihood ratio depends on whether the models are observationally identical or not. Two models are observationally identical if their probability densities are the same, when evaluated at the respective pseudo-true parameter values, i.e., $f_1(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{1*}) = f_2(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{2*})$ for almost all $(\mathbf{y}, \mathbf{x})$, where the pseudo-true parameter values, $\boldsymbol{\theta}_{k*}$, corresponds to the parameter value that minimizes the Kullback-Leibler distance from the true model (Sawa 1978).

Explicitly, Vuong showed (Theorem 3.3 on page 313) that under very general conditions:

1. If $f_1(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{1*}) = f_2(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{2*})$, then $2\, LR_{12}(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2)$ converges in distribution to a weighted sum of chi-square distributions.

2. If $f_1(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{1*}) \neq f_2(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{2*})$, then

$$\frac{1}{\sqrt{n}} \left(LR_{12}(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2) - E^0 \left[\log \frac{f_1(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{1*})}{f_2(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{2*})} \right] \right) \rightarrow^d \; N(0, \sigma_{12.12})$$

Because of this interesting asymptotic behavior Vuong had to proposed 3 distinct model selection tests: one for strictly non-nested models, that are always not observationally identical; another for overlapping models that might or might not be observationally identical; and a third for nested models, that are always observationally identical. (Nested models are always observationally identical because the nested model cannot be better than the full model and both models are equally close to the true model if and only if they are the same.)

In our applications, models M_1, M_2 and M_3 are not nested on each other, but are nested on models M_4^a, M_4^c and M_4^b, respectively (Figure 1 in the main text). Hence, our model selection tests consider pairs of models that are either non-nested or nested. In the Methods section we presented Vuong's test for not observationally identical models, that is suitable for the comparison of strictly non-nested models ($M_1 \times M_2$, $M_1 \times M_3$ and $M_2 \times M_3$).

We point out, however, that even though we perform model selection tests between nested models ($M_1 \times M_4$, $M_2 \times M_4$ and $M_3 \times M_4$) we don't need to use Vuong's test for nested models because our test statistics are based on penalized log-likelihoods instead of log-likelihoods, and our penalized models are not observationally identical for nested models too. In other words, even though $f_1(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{1*}) = f_4(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{4*})$ when model 1 is nested in model 4, we have that $f_1(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{1*}) - p_1 \neq f_4(\mathbf{y} \mid \mathbf{x}\,;\, \boldsymbol{\theta}_{4*}) - p_4$ since the penalty p_1 is smaller than p_4. Therefore, we can simply use Vuong's test for not observationally identical models in this case too.

On a technical note, we point out that Vuong's Theorem 3.3 still holds when we replace the log-likelihood ratio by the penalized log-likelihood ratio. The demonstration mimics Vuong's original proof presented on page 327. We just need to replace the log-likelihoods by penalized log-likelihoods in the Taylor expansion of the log-likelihoods around the maximum likelihood estimates.