Supplement B: Proof of Corollaries 2.5 and 2.6. Proof of Corollary 2.5. It is easy to see that Assumption 2.2 holds with $C_1 = C_2 = 1$ and Assumption 2.8 holds with $\delta = 0$ and $C_3 = 1$. We still use the method of proof of Theorem 2.2. Note that the following equation holds,

$$
- \log \left( \frac{p(\gamma|Z)}{p(\gamma^0|Z)} \right) = - \log \left( \frac{p(\gamma)}{p(\gamma^0)} \right) + \frac{1}{2} \log \left( \frac{\det(W_\gamma)}{\det(W_{\gamma^0})} \right) + \frac{n + \nu}{2} \log \left( \frac{1 + y^T(I_n - X_\gamma U_{\gamma^{-1}} X_\gamma^T)y}{1 + y^T(I_n - X_{\gamma^0} U_{\gamma^0^{-1}} X_{\gamma^0}^T)y} \right),
$$

(5.18)

Denote the above summands to be $J_1, J_2, J_3$. Following Assumption 2.1, $J_1$ is bounded below. By Lemma 5.2, if $\gamma \in S_1$, then $J_2 \geq 2^1(|\gamma| - s_n)\log(1 + n\phi_n)$; if $\gamma \in S_2$, then $J_2 \geq -2^{-1}s_n\log(1 + n\phi_n)$.

Now we approximate $J_3$. Note that $E \left( (\beta_{\gamma^0}^0)^T X_{\gamma^0}^T \epsilon \gamma^0 X_{\gamma^0}^T \epsilon \right) = \sigma_0^2 nk_n$, and hence $|E(\beta_{\gamma^0}^0)^T X_{\gamma^0}^T \epsilon|/(nk_n) = O_p(1/(\sqrt{n\psi_n})) = o_p(1)$. By a direct calculation and $s_n = o(n)$ (Assumption 2.9(i)), and the fact that $U_{\gamma} = (n + \phi_n^{-1})I_{|\gamma|}$ (by orthogonality of $X$), we can show that

$$
y^T(I_n - X_{\gamma^0} U_{\gamma^0^{-1}} X_{\gamma^0}^T)y = y^T(I_n - X_{\gamma^0} X_{\gamma^0}^T/(n + \phi_n^{-1}))y
$$

$$
= \frac{n\phi_n^{-1}}{n + \phi_n^{-1}} ||\beta_{\gamma^0}^0||_2^2 + \frac{2\phi_n^{-1}}{n + \phi_n^{-1}} (\beta_{\gamma^0}^0)^T X_{\gamma^0} \epsilon + \epsilon^T(I_n - X_{\gamma^0} X_{\gamma^0}^T/(n + \phi_n^{-1}))\epsilon
$$

$$
= \frac{n\phi_n^{-1}}{n + \phi_n^{-1}} k_n \left( 1 + \frac{2(\beta_{\gamma^0}^0)^T X_{\gamma^0} \epsilon}{nk_n} \right) + n\sigma_0^2 (1 + o_p(1))
$$

(5.19)

Let $\nu_{\gamma} = (I_n - P_{\gamma}) X_{\gamma^0} \gamma_{\gamma^0} \gamma_{\gamma}$. By the orthogonality of $X$, $\nu_{\gamma} = X_{\gamma^0} \gamma_{\gamma^0} \gamma_{\gamma}$. Then

$$
\max_{\gamma \in S_2} \frac{|\nu_{\gamma}^T \epsilon|}{||\nu_{\gamma}||_2} = \max_{\gamma \in S_2} \frac{|(\beta_{\gamma^0}^0)^T X_{\gamma^0} \gamma_{\gamma} \epsilon|}{||X_{\gamma^0} \beta_{\gamma^0}^0 \gamma_{\gamma}||_2} \leq \max_{\gamma \in \gamma^0} \frac{|(\beta_{\gamma}^0)^T X_{\gamma} \epsilon|}{||X_{\gamma} \beta_{\gamma}^0||_2}.
$$

Therefore, by Bonferroni’s inequality, we get that

$$
pr \left( \max_{\gamma \in S_2} \frac{|\nu_{\gamma}^T \epsilon|}{||\nu_{\gamma}||_2} > t \right) \leq pr \left( \max_{\gamma \in \gamma^0} \frac{|(\beta_{\gamma}^0)^T X_{\gamma} \epsilon|}{||X_{\gamma} \beta_{\gamma}^0||_2} > t \right) \leq \sum_{\gamma \in \gamma^0} pr \left( \frac{|(\beta_{\gamma}^0)^T X_{\gamma} \epsilon|}{||X_{\gamma} \beta_{\gamma}^0||_2} > t \right) \leq C_0 2^s n \exp \left( - \frac{t^2}{2\sigma_0^2} \right).
$$

Letting $t = C_0 \sigma_0 \sqrt{2sn}$ for a sufficiently large $C$, we can show that $\max_{\gamma \in S_2} |\nu_{\gamma}^T \epsilon|/||\nu_{\gamma}||_2 = O_p(\sqrt{s_n})$.

By Assumption 2.9(i), $s_n = o(n\psi_n^2)$. Therefore, by a similar proof to (5.5), uniformly for $\gamma \in S_2$,

$$
y^T \left( I_n - X_{\gamma} U_{\gamma}^{-1} X_{\gamma}^T \right) y \geq y^T (I_n - P_{\gamma}) y \geq n\psi_n^2 (1 + o_p(1))
$$

(5.20)
Consequently, by (5.19) and (5.20), we can see that with large probability and uniformly for $\gamma \in S_2$,

$$J_3 \geq \frac{n + \nu}{2} \log \left( \frac{n \psi^2_n}{\sigma^2_0 a_n} \right).$$

When $\gamma \in S_1$,

$$\frac{1 + y^T(I_n - X_\gamma U^{-1}_\gamma X^T_\gamma)}{1 + y^T(I_n - X_\gamma U^{-1}_\gamma X^T_\gamma)} = 1 - \frac{y^T(X_\gamma X^T_\gamma - X_\gamma X^T_\gamma)y/(n + \phi_n^{-1})}{1 + y^T(I_n - X_\gamma X^T_\gamma/(n + \phi_n^{-1}))y} = 1 - \frac{\epsilon^T(X_\gamma X^T_\gamma - X_\gamma X^T_\gamma) \epsilon/(n + \phi_n^{-1})}{1 + \sigma^2_0 a_n(1 + o_p(1))}.$$

Following the proofs of Lemma 3 by Meinshausen and Yu (2009), with large probability, uniformly for $\gamma \in S_1$, $\epsilon^T(X_\gamma X^T_\gamma - X_\gamma X^T_\gamma) \epsilon \leq 2n \sigma^2(\lvert \gamma \rvert - s_n) \log p$. Note that $p \log p = o(a_n)$ (Assumption 2.9(ii)), and the inequality $\log(1 - x) \geq -(\alpha/2)x$ holds when $x \in (0,1 - 2/\alpha)$ for any $2 < \alpha < \alpha_0$. Thus, when $n$ is sufficiently large, with large probability, for any $\gamma \in S_1$,

$$J_3 = \frac{n + \nu}{2} \log \left( 1 - \frac{1}{n + \phi_n^{-1}} \frac{\epsilon^T(X_\gamma X^T_\gamma - X_\gamma X^T_\gamma) \epsilon}{1 + \sigma^2_0 a_n(1 + o_p(1))} \right) \geq \frac{n + \nu}{2} \log \left( 1 - \frac{2(\lvert \gamma \rvert - s_n) \log p}{a_n} \right) \geq -2^{-1} \left( \frac{n}{\alpha_0} \frac{(n + \nu)}{\alpha_0} \frac{\log p}{a_n} \right).$$

Then, $\max_{\gamma \neq \gamma^0} p(\gamma \mid Z) / p(\gamma^0 \mid Z) \rightarrow_p 0$ follows by $s_n = o \left( \frac{n \psi^2_n}{\sigma^2_0 a_n} \frac{\log n}{\log(1 + o(a_n))} \right)$ (Assumption 2.9(i)) and assumption $p^{\alpha_2(n + \nu)/a_n} = o(n \phi_n)$, and by applying arguments similar to (5.8) and (5.9). By $p = o \left( \frac{(n + \nu) \log n}{\alpha_0} \frac{\log p}{a_n} \right)$, $p^{2+\alpha_2(n + \nu)/a_n} = o(n \phi_n)$, and arguments similar to (5.10) and (5.11), it can be shown that $\sum_{\gamma \neq \gamma^0} p(\gamma \mid Z) / p(\gamma^0 \mid Z) \rightarrow_p 0$, and therefore $p(\gamma^0 \mid Z) \rightarrow_p 1$.

**Proof of Corollary 2.6.** (a) Assumptions 2.1 and 2.6 are easy to verify. Define $\alpha_n \sim \beta_n$ to mean that there exist positive constants $d_1$ and $d_2$ such that $d_1 < \alpha_n / \beta_n < d_2$ when $n$ goes to $\infty$. It is easy to see that $a_n \sim n^{1+\delta_1-\delta_2} (\log n)^2$ and $n \psi^2_n \sim n^{2+\delta_1} (\log n)^2$, therefore $n \psi^2_n \gg a_n$, so Assumption 2.9 can be verified by a direct calculation. Notice that $(n + \nu) / a_n \rightarrow 0$, $n^{\alpha_2(n + \nu)/a_n} = o(n \phi_n)$ holds for any $\alpha_0$. By Corollary 2.5, $\max_{\gamma \neq \gamma^0} p(\gamma \mid Z) / p(\gamma^0 \mid Z) \rightarrow_p 0$. When $1 < \delta_2 \leq \delta_1$, it is easy to see that $n = o \left( \frac{(n + \nu) \log n}{\alpha_0} \right)$ and $n^{2+\alpha_2(n + \nu)/a_n} = o(n \phi_n)$, so by Corollary 2.5, $p(\gamma^0 \mid Z) \rightarrow_p 1$.

Next, we assume $-1 < \delta_2 \leq 1$ and identify the limit of $p(\gamma^0 \mid Z)$. Without loss of generality, let $\beta_j^0 \neq 0$, $1 \leq j \leq s$, and $\beta_{s+1}^0 = \cdots = \beta_n^0 = 0$. For $s + 1 \leq j \leq n$, define $\gamma(j)$ to be an $n$-vector with 1s in the first
s positions and jth position, and zero in others, i.e., \( \gamma(j)_1 = \cdots = \gamma(j)_s = \gamma(j)_{j+1} = 1, \gamma(j)_i = 0 \) when \( i \neq 1, \cdots, s, j \). Clearly, each \( \gamma(j) \) corresponds to a model in \( S_1 \).

Denote \( \gamma = \gamma(j) \) for some \( s + 1 \leq j \leq n \), then \( \det(W_s) = (1 + n\phi_n)^{s+1} \) and \( \det(W_{\gamma,0}) = (1 + n\phi_n)^s \). Let \( J_1, J_2 \) and \( J_3 \) be defined as in (5.18). Consequently, \( J_2 = \frac{1}{2} \log(1 + n\phi_n) \). Using the representation (5.18) and the fact that \( J_1 = 0 \) and \( J_3 \) is almost surely non-positive when \( \gamma \in S_1 \), we have

\[
(5.21) \quad - \log \left( p(\gamma|Z)/p(\gamma^0|Z) \right) \leq \frac{1}{2} \log(1 + n\phi_n), \text{ a.s.,}
\]

and so \( p(\gamma|Z)/p(\gamma^0|Z) \geq (1 + n\phi_n)^{-1/2} \), a.s., which leads to

\[
(5.22) \quad \sum_{\gamma \neq \gamma^0} p(\gamma|Z)/p(\gamma^0|Z) \geq \sum_{j=s+1}^{n} p(\gamma(j)|Z)/p(\gamma^0|Z) \geq \frac{n-s}{(1 + n\phi_n)^{1/2}}, \text{ a.s.}
\]

When \(-1 < \delta_2 < 1\), it follows from (5.22) that \( \sum_{\gamma \neq \gamma^0} p(\gamma|Z)/p(\gamma^0|Z) \to \infty \), a.s. Therefore, \( p(\gamma^0|Z) \to 0 \), a.s. follows from relationship (2.7). When \( \delta_2 = 1 \), \( \frac{\sum_{j=s+1}^{n} p(\gamma(j)|Z)/p(\gamma^0|Z)}{(1 + n\phi_n)^{1/2}} \) converges to some positive constant \( c \). It thus follows from (5.22) that \( \limsup_n p(\gamma^0|Z) \leq c_0 := 1/(1 + c), \) a.s.

(b) Using the relationship (2.7), it is sufficient to prove \( p(\emptyset|Z)/p(\gamma^0|Z) \to \infty \). Let \( J_1, J_2, J_3 \) be defined as in the representation (5.18). Then \( J_1 = 0 \) follows from assumption. It is easy to see that \( J_2 = -\frac{3}{2} \log(1 + n\phi_n) \). Next, we approximate \( J_3 \). Since \( n^{\log n} = O(\phi_n), a_n \sim n \). Thus, by the proof of (5.19), \( y^T \left( I_n - X_{\gamma,0} U_{\gamma,0}^{-1} X_{\gamma,0}^T \right) y = \sigma_0^2 a_n(1 + o_p(1)) \sim \sigma_0^2 n(1 + o_p(1)) \). Since \( \epsilon^T X_{\gamma,0} \beta_{\gamma,0}^0 = O_p \left( |nk_n|^{1/2} \right) \), \( \epsilon^T X_{\gamma,0} X_{\gamma,0}^T \epsilon = O_p(n) \) and \( X_{\gamma,0}^T X_{\gamma,0} = nI_s \), and by the fact that \( nk_n \geq s\nu \psi_n^2 \to \infty \), we have

\[
\begin{align*}
\mathbf{y}^T X_{\gamma,0} X_{\gamma,0}^T \mathbf{y} & = \left( X_{\gamma,0} \beta_{\gamma,0}^0 + \epsilon \right)^T X_{\gamma,0} X_{\gamma,0}^T \left( X_{\gamma,0} \beta_{\gamma,0}^0 + \epsilon \right) \\
& = n^2 \| \beta_{\gamma,0}^0 \|_2^2 + 2n \epsilon^T X_{\gamma,0} \beta_{\gamma,0}^0 + \epsilon^T X_{\gamma,0} X_{\gamma,0}^T \epsilon \\
& = n^2 k_n + O_p(n^{(nk_n)^{1/2}}) + O_p(n) \\
& = n^2 k_n (1 + o_p(1)),
\end{align*}
\]

and so it is easy to see that

\[
J_3 = \frac{n + \nu}{2} \log \left( \frac{1 + \frac{1}{n + \phi_n^{-1}}} {1 + \frac{\mathbf{y}^T X_{\gamma,0} X_{\gamma,0}^T \mathbf{y}} {1 + \mathbf{y}^T \left( I_n - X_{\gamma,0} U_{\gamma,0}^{-1} X_{\gamma,0}^T \right) \mathbf{y}}} \right) = O_p \left( \frac{n + \nu}{2} \log(1 + k_n) \right).
\]
By a direct calculation, it is not hard to verify that $(n + \nu) \log(1 + k_n) \ll \log(1 + n\phi_n)$, so

\[- \log \left( \frac{p(\emptyset|Z)}{p(\gamma^0|Z)} \right) = -\frac{s}{2} \log (1 + n\phi_n) + O_p \left( \frac{n + \nu}{2} \log (1 + k_n) \right) \]

\[= -\frac{s}{2} \log (1 + n\phi_n) (1 + o_p(1)),\]

which leads to $p(\emptyset|Z)/p(\gamma^0|Z) \to_p \infty$. Thus, $p(\gamma^0|Z) \to_p 0$ follows immediately from (2.7).

(c) The proof can be finished by constructing a sequence of models containing the true model using the approach in part (a) and by arguments (5.21) and (5.22). \qed