unidades dentro daquele nível – esse tipo de modelo é chamado de efeito misto justamente por considerar a existência de efeito fixo e de efeito aleatório.

Inicialmente pensamos em fazer o agrupamento das observações por UF usando os Partidos Políticos como um terceiro nível, dentro da UF. Entretanto, muitos partidos apresentaram somente um candidato em algumas UFs, o que inviabiliza a sua inclusão como um nível de análise à parte no método hierárquico. Por esse motivo, os partidos políticos entraram no modelo como co-variável e não como um nível de análise.

Após a definição das variáveis de interesse, a execução da coleta dos dados e o estudo das variáveis, procedeu-se ao processo de ajuste do modelo – descrito adiante.

Conforme já vimos, a variável resposta Y_{ij} é o resultado eleitoral Y do candidato i na UF j. Y_{ij} pode assumir dois valores:

$$Y_{ij} = \begin{cases} 0, \textit{ se não eleito} \\ 1, \textit{ se eleito} \end{cases} \tag{1}$$

Assim, a probabilidade de um candidato ser eleito é:

$$\Pr(Y_{ij} = 1) = \phi_{ij} \tag{2}$$

Aplicando a função de ligação logit à probabilidade de eleição ϕ_{ij}, temos o log das chances de eleição η_{ij}:

$$\eta_{ij} = \log\left(\frac{\phi_{ij}}{1-\phi_{ij}}\right) \tag{3}$$

Tentamos inicialmente ajustar um modelo de η_{ij} que incluísse todas as variáveis explicativas que pretendíamos estudar, para em seguida ir retirando, uma a uma, as menos significativas estatisticamente. Contudo, o grande número de variáveis fez com que a matriz do modelo ficasse quase singular, provavelmente por multicolinearidade entre as variáveis. Optamos então por iniciar o ajuste sem os partidos, que sozinhos representavam 28 variáveis (uma variável binomial para cada partido, ficando de fora o PV, usado como referência). O modelo ajustado inicialmente para η_{ij} pode ser descrito formalmente como:

$$\begin{aligned}
\eta_{ij} &= \beta_{0j} + \beta_{1j}(SEXO1)_{ij} + \beta_{2j}(PRIMCOUM)_{ij} + \beta_{3j}(SEGCOUMA)_{ij} + \beta_{4j}(SUPCOMPL)_{ij} \\
&+ \beta_{5j}(CAS)_{ij} + \beta_{6j}(Idade)_{ij} + \beta_{7j}(REELEICA)_{ij} \\
\beta_{0j} &= \gamma_0 + \alpha_1(CANDPVAG)_j + \alpha_2(TAMDIST)_j + \alpha_3(POPURBAN)_j + \alpha_4(DENSIDAD)_j \\
&+ \alpha_5(EXPECT)_j + \alpha_6(TAXAALFA)_j + \alpha_7(PIB_PER)_j + \alpha_8(ELEITORA)_j + \alpha_9(RZSEXELE)_j \\
&+ \alpha_{10}(ELEITPCA)_j + u_{0j} \\
\beta_{1j} &= \gamma_1 + u_{1j} \qquad (4) \\
\beta_{2j} &= \gamma_2 + u_{2j} \\
\beta_{3j} &= \gamma_3 + u_{3j} \\
\beta_{4j} &= \gamma_4 + u_{4j} \\
\beta_{5j} &= \gamma_5 + u_{5j} \\
\beta_{6j} &= \gamma_6 + u_{6j} \\
\beta_{7j} &= \gamma_7 + u_{7j}
\end{aligned}$$

No modelo expresso na equação 4, i é o índice dos candidatos $\{i \in \mathrm{N} / 1 \geq i \geq 4946\}$, j o índice das UFs $\{j \in \mathrm{N} / 1 \geq j \geq 27\}$. Sendo s o índice das variáveis do nível 2 (UFs) $\{s \in \mathrm{N} / 1 \geq s \geq 10\}$ e q o índice das variáveis do nível 1 (candidatos) $\{q \in \mathrm{N} / 1 \geq q \geq 7\}$, α_S