$$\begin{aligned}\log(\text{incwage}_i) = \beta_0 &+ \beta_1 \cdot \text{STEMType}_i + \beta_2 \cdot \text{Citizenship}_i + \beta_3 \cdot (\text{STEMType}_i \times \text{Citizenship}_i) \\ &+ \beta_4 \cdot \text{AgeGroup}_i + \beta_5 \cdot \text{Sex}_i + \beta_6 \cdot \text{Education}_i + \beta_7 \cdot \text{Race}_i \\ &+ \beta_8 \cdot \text{MaritalStatus}_i + \beta_9 \cdot \text{ClassOfWork}_i + \beta_{10} \cdot \text{IndustrySector}_i \\ &+ \beta_{11} \cdot \text{WorkTime}_i + \beta_{12} \cdot \text{Disability}_i + \varepsilon_i \end{aligned} \tag{1}$$

This OLS regression model estimates the association between working in a STEM occupation and log-transformed wage income. The model controls for key demographic and labor market characteristics, including age, gender, education, race, marital status, employment type, industry, and work hours. The specification also includes an interaction between STEM type and citizenship to capture differential returns to STEM work across citizenship groups. Standard errors are robust (HC3), and results should be interpreted as correlations, not causal effects.