



Additional Figure 2: Levels defined in terms of the minimum distance from the root (node 1) lead to inconsistent predictions. The small numbers close to nodes correspond to the \hat{y}_i scores of the flat predictions. The Hierarchical top-down scores obtained respectively by crossing the levels according to the minimum and the maximum distance from the root are shown in the bottom-left. Scores in boldface represent inconsistent predictions.