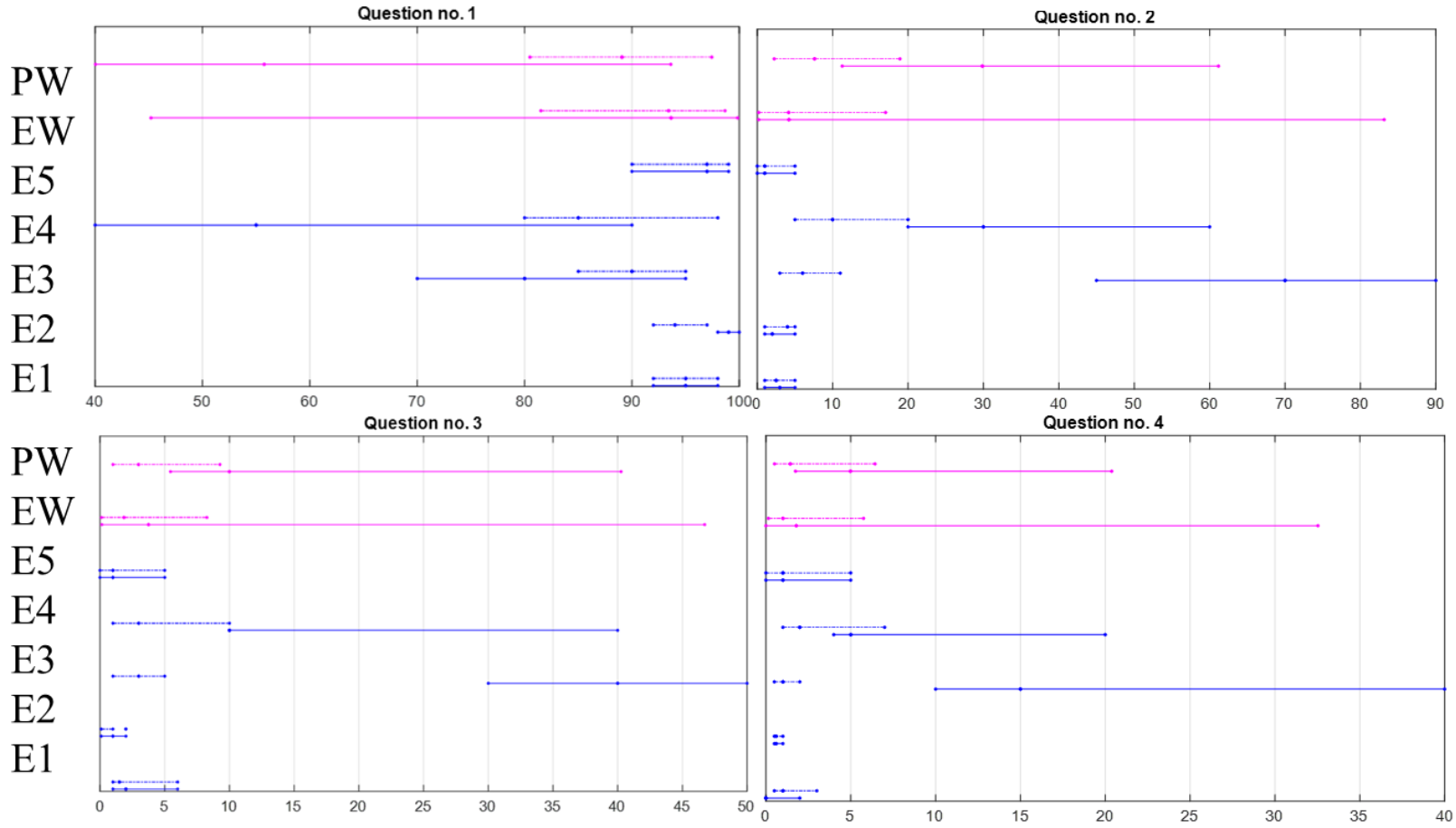
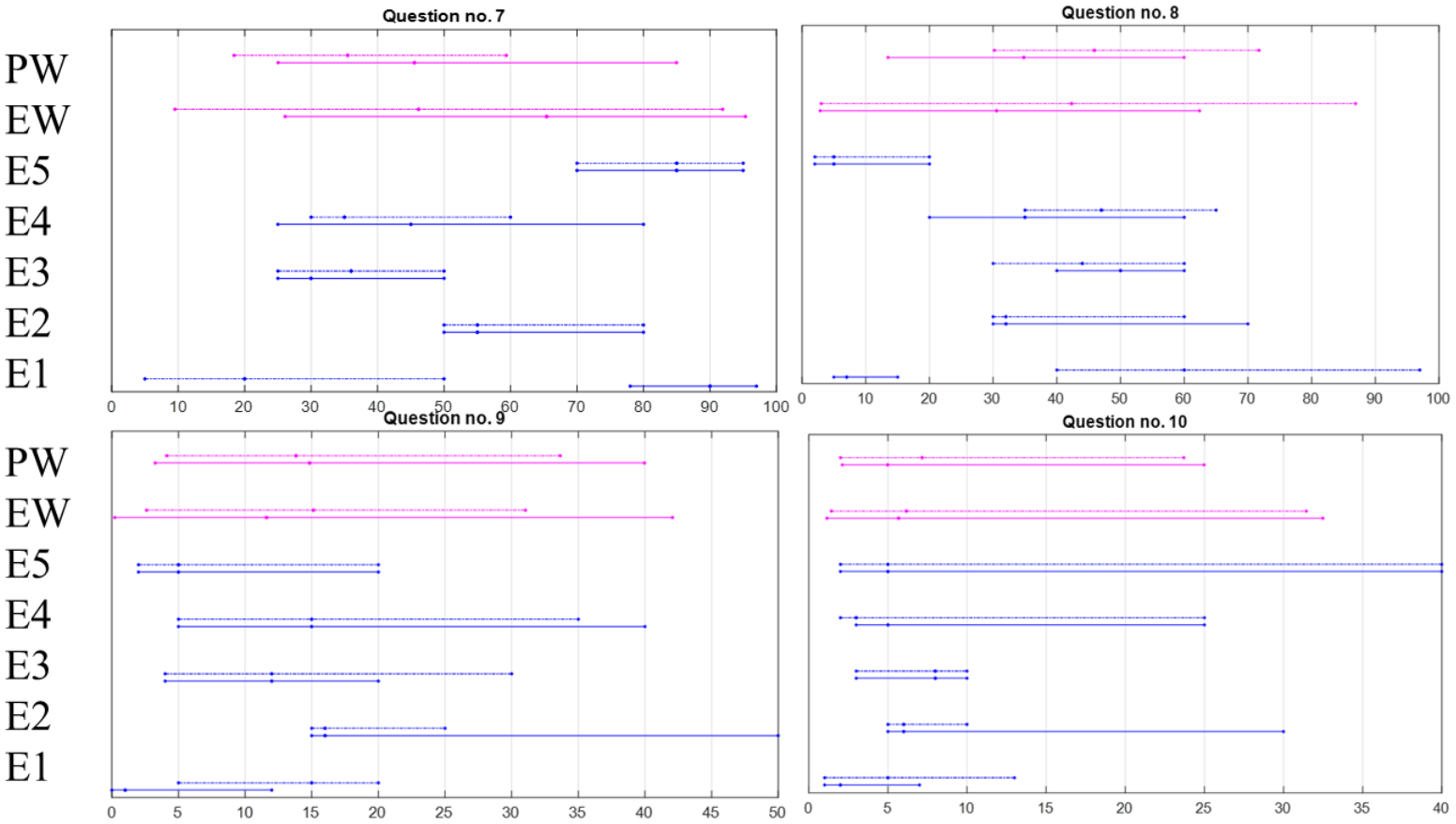
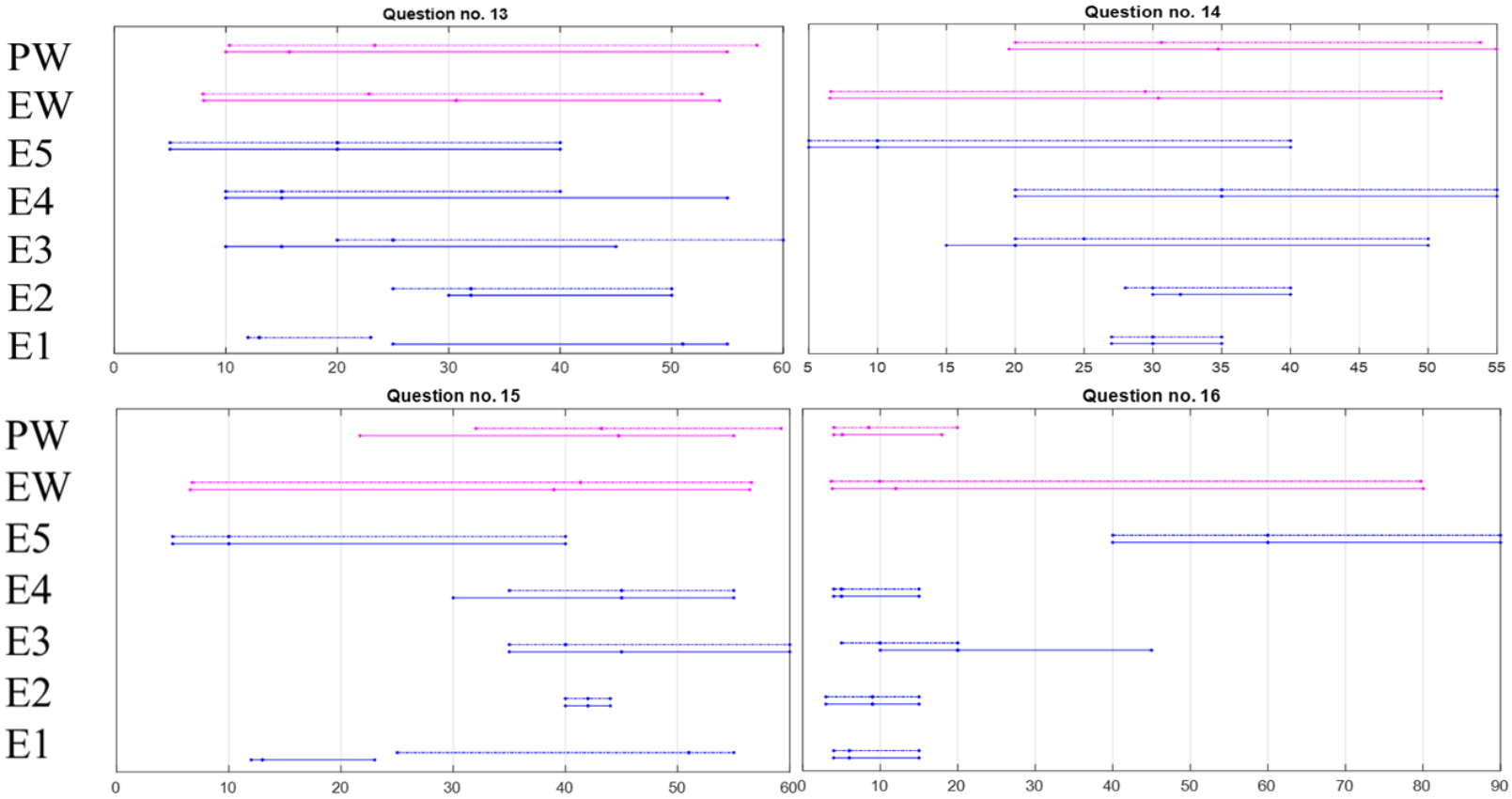


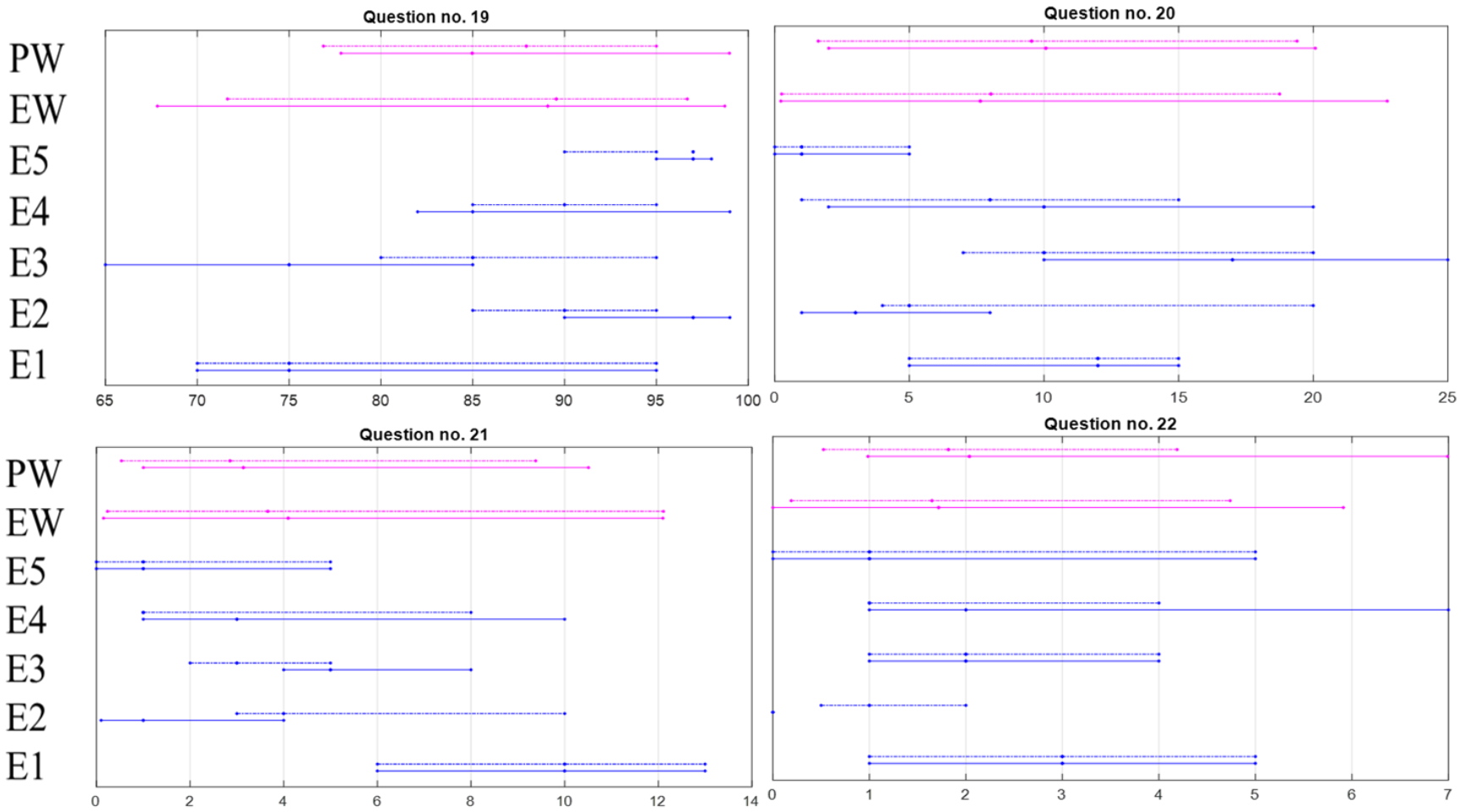
Supplementary Figure 1 Elicitation Q's1-4: Performance weighted (PW) and equal-weighted (EW) pooled estimates converge in Round 2. Experts 3 (E3) & 4 (E4) (numbering from the bottom) redistribute their 100 people between categories in Round 2, with more allocated to 'food secure' (Q1 increases) and other categories (Q2-4) reduced accordingly. Their uncertainty bands, the distance between the lowest plausible value and highest plausible value, are also reduced substantially after discussion.



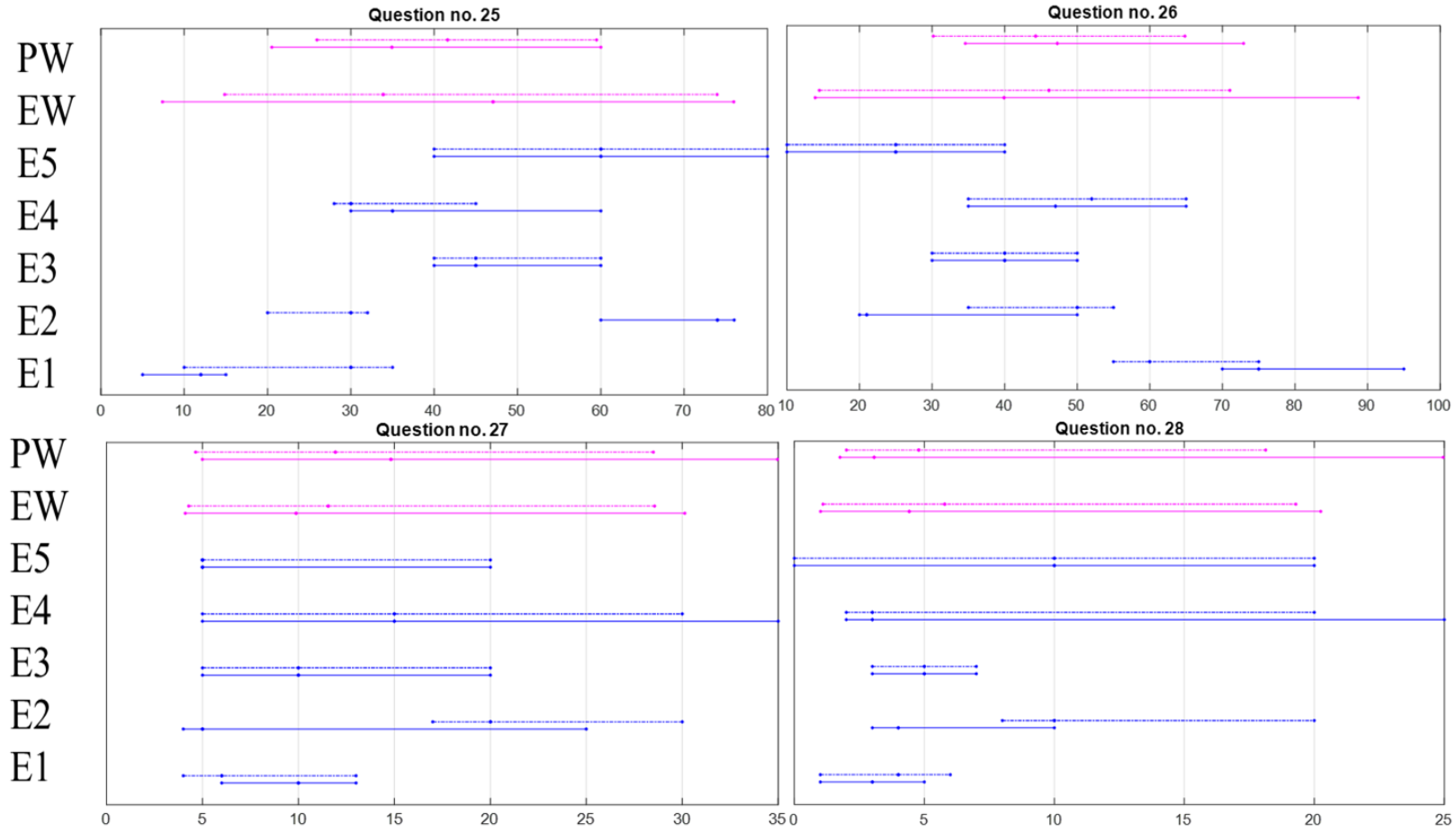
Supplementary Figure 2 Elicitation Q's7-10: Performance weighted (PW) & equal weighted (EW) pooled estimates converge in Round 2. Expert 1(E1) redistributes their 100 people in Round 2 between categories, reducing the 'Food Secure' (Q7) and increasing the other categories. Expert 4(E4) redistributes some from 'Food Secure' (Q7) to 'Moderately Food Secure' (Q8) in Round 2 and Expert 3 (E3) makes the opposite redistribution, converging on similar values for Round 2. As in the calibration questions, Expert 1 changes most between Round 1 & Round 2, and Expert 5 (E5) does not change in and uses very wide uncertainty bands (the difference between the highest plausible and lowest plausible values), especially for the 'Low Food Secure' and Very Low Food Secure' categories here (Q9 &10).



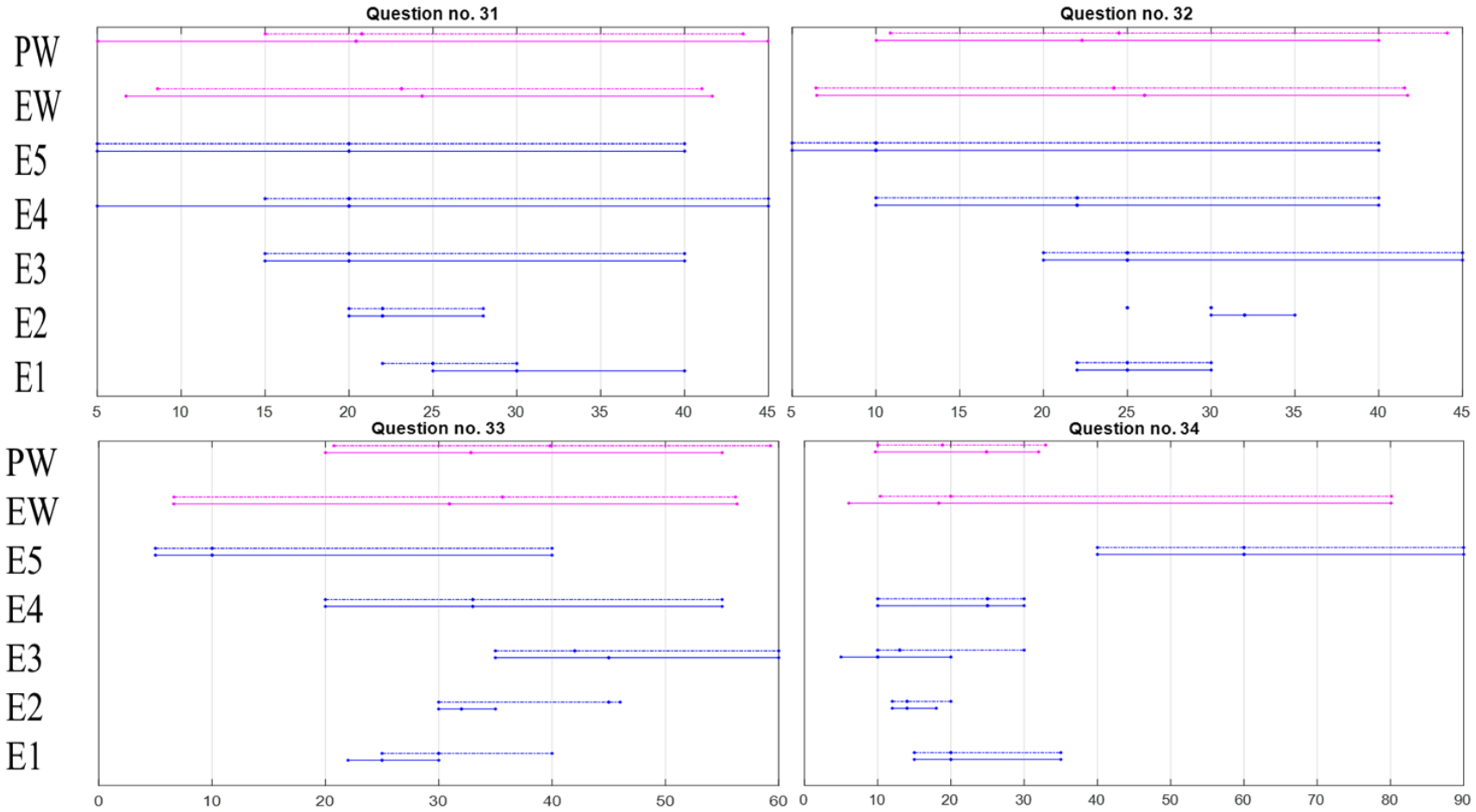
Supplementary Figure 3 Elicitation Q's13-16: Performance weighted (PW) & equal weighted (EW) converge in Round 2. Expert 1(E1) redistributes people from the 'Food Secure' category (Q13) to the 'Low Food Secure' category (Q15). Expert 3(E3) redistributes people from 'Very Low Food Secure' to 'Food Secure' and Moderately Food Secure'. The other experts do not change and Expert 5(E5) uses very wide uncertainty bands (the difference between the highest plausible and lowest plausible values) throughout.



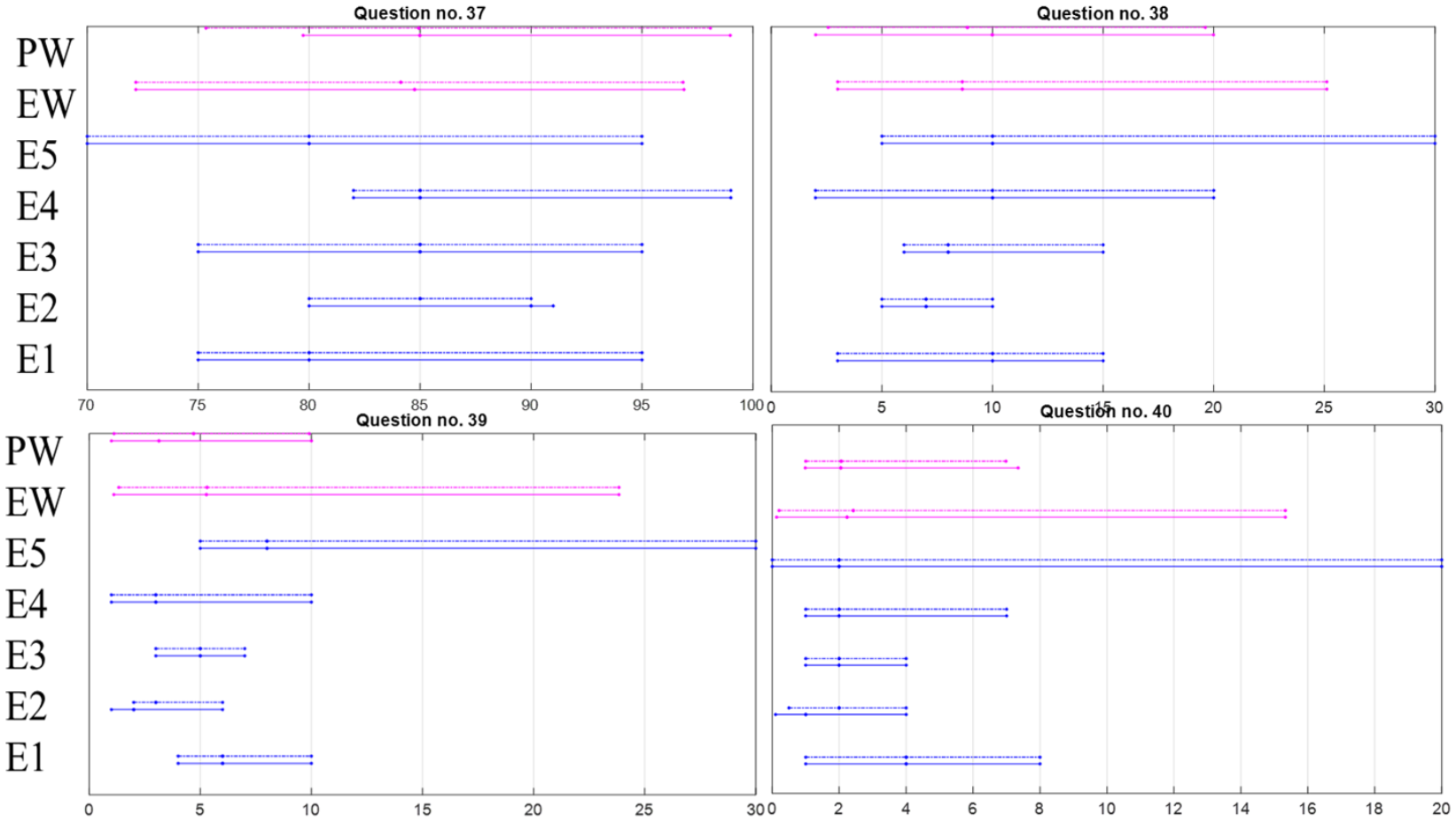
Supplementary Figure 4 Elicitation Q's19-22: Performance weighted & equal weighted converge in Round 2. Expert 2 increased the 'Very Low Food Secure' from 0, increasing also the 'Low Food Secure' and 'Moderately Food secure' by reducing number in 'Food Secure'. Expert 3 increased the numbers in 'Food Secure' by reducing other categories, converging towards the values given by other experts.



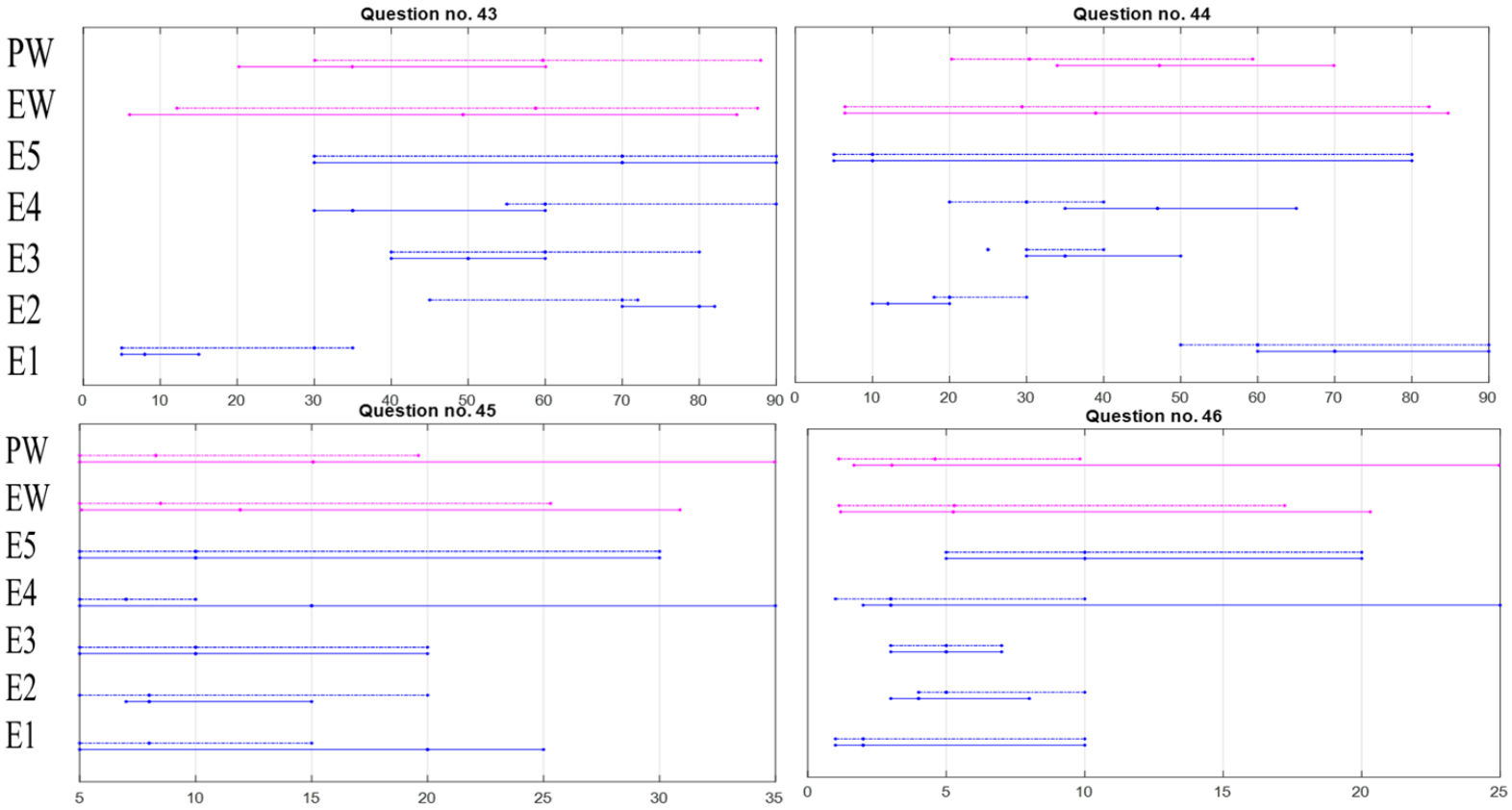
Supplementary Figure 5 Elicitation Q's25-28: Performance weighted (PW) & equal weighted (EW) converged in Round 2. Expert 4 (E4) made a small shift from 'Food Secure' to 'Moderately Food Secure' in Round 2. Expert 2 (E2) made the most Round 2 changes, reducing 'Food Secure' and increasing all other categories. Expert 1(E1) increased 'Food Secure' & 'Moderately Food Secure', reduced 'Low Food Secure' and increased 'Very Low Food Secure', mostly aligning more closely with other experts.



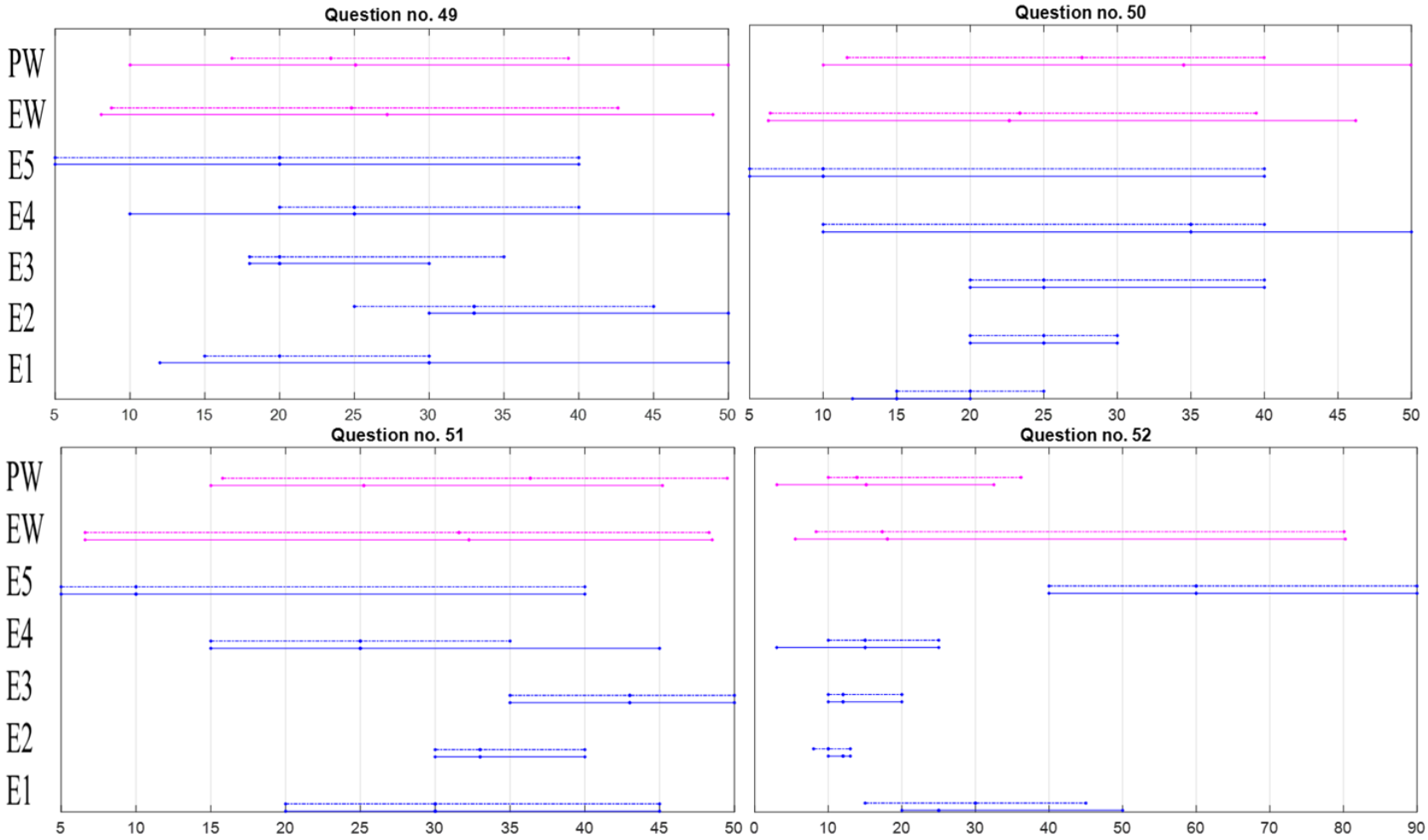
Supplementary Figure 6 Elicitation Q's31-34: Performance weighted (PW) & equal weighted EW) converge in Round 2. Expert 1 (E1) made minor adjustments between categories, but no major shifts to report.



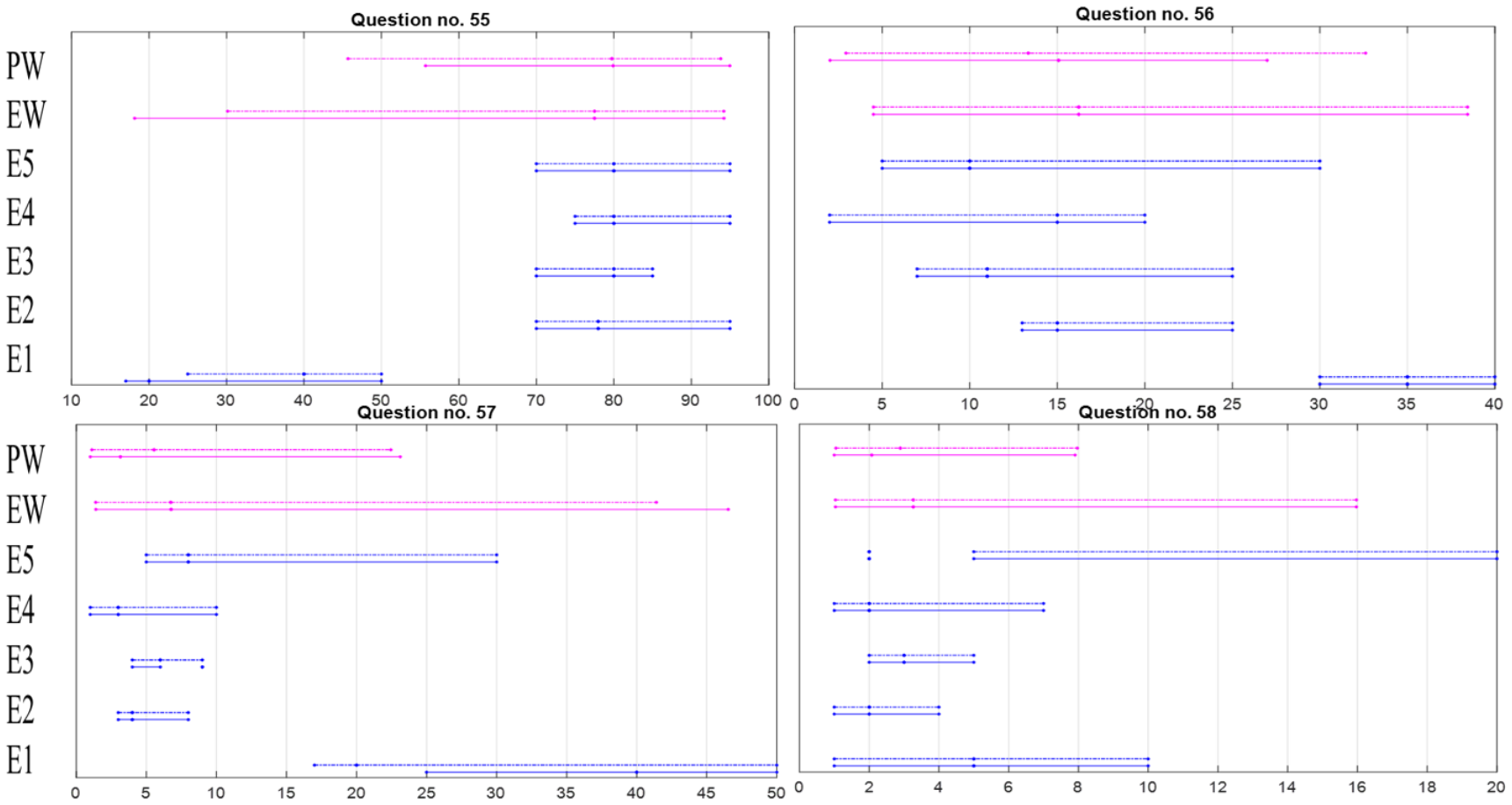
Supplementary Figure 7 Elicitation Q's37-40: Limited difference between Round 1 and Round 2.



Supplementary Figure 8 Elicitation Q's43-46: Performance weighted (PW) & equal weighted (EW) pooled estimates converge in Round 2. Expert 5 (E5) made no changes and used very wide confidence bands. Experts 1, 3 & 4 (E1,E3,E4) increase the numbers in 'Food Secure' and reduced 'Moderately Food Secure' categories. Experts 1 & 4 (E1 & E4) also reduced 'Low Food Secure'. Expert 2(E2) reduced 'Food Secure' and increased 'Moderately Food Secure' and 'Very Low Food Secure'



Supplementary Figure 9 Elicitation Q's49-52: Performance weighted (PW) & equal weighted (EW) converged in Round 2. Expert 1(E1) reduced the numbers in' Food Secure' and increased 'Moderately Food Secure'. Other experts made minor changes to uncertainty intervals (the difference between the highest plausible and lowest plausible values), but not to best estimates.



Supplementary Figure 10 Elicitation Q's55-58: Performance weighted (PW) & equal weighted (EW) converged in Round 2. No expert made significant changes between Round 1 & Round 2.