*Supplementary Table 1*: CHARLS 6-item validation scale

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Domain* | *Operationalized as* | *Operational Limitations* | | *Interpretation* |
| ***Number of NAs (%)***  ***Imputation Method*** | ***Limitations*** |
| *Population density* | Scored on a scale of 0 ~ 10 based on cut-off values obtained from CHNS urbanicity scale | 27 (5.9%)  Average county-level values |  | Assumption: |
| *Economic activity* | The scores are divided into upper and lower groups at the median. They are then scored on a scale of 0 to 10 at 0.5-point intervals by dividing the upper and lower groups into 11 and 10 bins of equal size respectively. | 2 (0.4%)  Average county-level values |  | Assumption: |
| *Markets* |  | 6 (1.3%)  Replace with 0 | No justifiable means of imputation was available. All missing or invalid entries were replaced with 0. The degree to which this imputation skews the estimated distribution is unknown. | Assumption: |
|  | 5 (1.1%)  Replace with 0 |
|  | 4 (0.8%)  Replace with 0 |
|  | |

*Supplementary Table 1 Continued*: CHARLS 6-item validation scale

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Domain* | *Operationalized as* | *Operational Limitations* | | Interpretation |
| ***Number of NAs (%)***  ***Imputation Method*** | **Limitations** |
| *Transportation* |  | 1 (0.2%)  Fill with 0 | No justifiable means of imputation was available. All missing or invalid entries were replaced with 0. The degree to which this imputation skews the estimated distribution is unknown. | Assumption: |
|  | 23 (5.1%)  Fill with 0 |
|  | 9 (2.0%)  Fill with 0 |
|  | |

*Supplementary Table 1 Continued*: CHARLS 6-item validation scale

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Domain* | *Operationalized as* | *Operational Limitations* | | *Interpretation* |
| ***Number of NAs (%)***  ***Imputation Method*** | ***Limitations*** |
| *Housing* |  | 2 (0.4%)  Average county-level values | A total of 71 communities had missing values for the proportion of households using surface or ground water. These were replaced with 0. Furthermore, 8 had missing values for tap water usage. These were imputed using average county-level values. | Assumption: |
|  | 71 (15.6%)  Replace with 0  Final score:  8 (1.7%)  Average county-level values |
|  | 68 (15.0%)  Replace with 0  Final score:  9 (2.0%)  Average county-level values | A total of 68 communities had missing values for the proportion of households using LPG, natural gas, coal, hay, marsh gas, or others as cooking fuel. These were replaced with 0. Furthermore, 9 communities had for all type of fuel use. These were imputed using average county-level values. |
|  | | |

*Supplementary Table 1 Continued*: CHARLS 6-item validation scale

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Domain* | *Operationalized as* | *Operational Limitations* | | *Interpretation* |
| ***Number of NAs (%)***  ***Imputation Method*** | ***Limitations*** |
| *Diversity* | The variance scores are divided into upper and lower groups at the median. They are then scored on a scale of 0 to 10 at 0.5-point intervals by dividing the upper and lower groups into 11 and 10 bins of equal size respectively. | 9 (2.0%)  Average county-level values  Average province-level values |  | Assumption: |
| The variance scores are divided into upper and lower groups at the median. They are then scored on a scale of 0 to 10 at 0.5-point intervals by dividing the upper and lower groups into 11 and 10 bins of equal size respectively. | 2 (0.4%)  Not Imputed |  |
|  | |  |