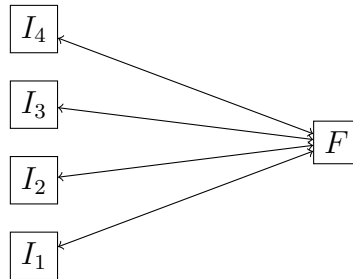


Supplementary Information

Keywords: Slums, Factor Analysis, Deprivation Index, Gray-Level Co-Occurrence Matrix, Mexico City

1. Calculation of weights using Factor Analysis

In factor analysis, we assume a particular causal structure with respect to the relationship between the unobserved “factor” (i.e., slum severity) and the “indicators” that we have at hand (i.e., overcrowding, sanitation, structural quality and safe water) as shown below.



Here, some unmeasured factor, F (i.e., slum severity) is correlated with and helps to determine the amount of variation in the indicator variables (i.e., overcrowding, sanitation, structural quality and safe water) I_1 through I_4 . In other words, as F goes up or down, the indicator values all move up and down. The amount of correlation between F and I_1 through I_4 will determine the variation in the I variable as F changes. It has been shown mathematically that if F (i.e., slum severity) is correlated to, say, I_1 (i.e., overcrowding)

and I_2 (i.e., sanitation), then I_1 and I_2 will be correlated with one another, even if there is no intrinsic relationship between I_1 and I_2 other than their common relationship to F . Therefore, if F can cause correlation between the indicator variables, then we could use the pattern of correlation between the indicators (which can be ascertained, because we have data on the indicators from the INEGI Mexico) to determine the correlation coefficient between F and the indicators (I).

Therefore, factor analysis allows us to infer the unknown factor (i.e., slum severity) from the pattern of correlation between the indicators that we can measure. In other words, if one of the indicators (i.e., overcrowding, sanitation, structural quality and safe water) was completely determined by slum severity, then it would be a perfect proxy for slum severity index. The weight of that particular variable would be 1 and we could have ignored other variables. In factor analysis, this measure, also known as the communality, determines the amount of variation in the indicators that is due to the common factor and is a good proxy for weighting each indicator. Since communality is completely calculated from the input data; the weights are automatically derived and unique for each dataset which enables a context-based measurement of shelter deprivation in slums.