Strengthening Social Protection for Pandemic Responses

Guiding Poverty Reduction

A flexible modelling approach to nowcasting and forecasting Arab multidimensional poverty

UNITED NATIONS
الاسكوا
ESCWA
Shared Prosperity Dignified Life

UNITED NATIONS
الاسكوا
ESCWA

Shared Prosperity **Dignified Life**

VISION
ESCWA, an innovative catalyst for a stable, just and flourishing Arab region

MISSION
Committed to the 2030 Agenda, ESCWA's passionate team produces innovative knowledge, fosters regional consensus and delivers transformational policy advice. Together, we work for a sustainable future for all.

E/ESCWA/CL3.SEP/2021/TP.1

Economic and Social Commission for Western Asia

A flexible modelling approach to nowcasting and forecasting Arab multidimensional poverty

United Nations
Beirut

This paper was prepared by Paul Makdissi (Department of Economics, University of Ottawa, Canada; Email: paul.makdissi@uottawa.ca). The author would like to thank Khalid Abu-Ismail, Sama El Hage Sleiman, Vladimir Hlasny, Mohamad Khaled and Myra Yazbeck for their useful input for this paper.

United Nations publication issued by ESCWA, United Nations House, Riad El Solh Square,
P.O. Box: 11-8575, Beirut, Lebanon.

Website: www.unescwa.org.

Contents

List of tables

List of figures

Introduction

Quantifying inequality and poverty in the Arab region is a challenging task, particularly as statistical capacity in the Arab region remains weak (Atamanov and others, 2020). Few countries in the region monitor changes in the incomes and expenditures of their populations. In Arab countries that do conduct such surveys, data tends to be gathered and analysed on an irregular basis. Analysts must therefore adopt alternative approaches in order to assess inequality and poverty in the region. One alternative approach is to look at multidimensional poverty in the context of an MPI framework (Alkire and Foster, 2011; Alkire and Santos, 2014). That approach relies on the use of data collated in demographic health surveys. The Economic and Social Commission for Western Asia (ESCWA) has adapted the MPI approach to take into consideration the particular nature of Arab multidimensional poverty. (Economic and Social Council for Western Asia, 2020) Unfortunately, the surveys required to compute multidimensional poverty indices are conducted on an irregular basis in the Arab region. Given the limited data available, steps must sometimes be taken to nowcast poverty indices on the basis of two non-consecutive surveys and make forecasts beyond the most recent survey conducted.

Despite the broad assumptions made, nowcasting and forecasting approaches to computing poverty indices remain an essential exercise from a policy perspective. Even if there is uncertainty on the exact magnitude of an index, capturing at least the direction of change is essential if policymakers are to formulate successful evidence-based policies. A number of authors have proposed models to nowcast deprivation in specific dimensions such as health (Klassen and Lange, 2012). However, despite one or two notable exceptions (Ram, 2020; Alkire and others, 2020), little literature has been published to date on proposed models to nowcast multidimensional poverty.

This paper sets forth a modelling approach for nowcasting the Arab multidimensional poverty index. Our approach complements the work conducted by Alkire and others (2020) in that we propose a dynamic model of multidimensional poverty. Our approach differs, however, in that we model the change in each indicator separately. This analytical choice allows for more flexibility in terms of the modelling assumptions made to estimate the counterfactual change in the univariate distribution of each indicator. We aggregate the overall change by assuming the stability of the copula through time. We illustrate our methodology with reference to the 2011 and 2018 UNICEF Multiple Indicator Cluster Surveys for Iraq.

Section II of the present paper sets out a theoretical framework for modelling the Arab multidimensional poverty index. In Section III of the paper, we illustrate our theoretical framework using simple logistic models and by looking at the potential impact of exogenous shocks, such as the COVID-19 pandemic, to capture the resulting dynamics of deprivation in each indicator. In Section IV, we present the results of our simulation. The paper concludes by suggesting areas for future research and policy modelling.

1. Modelling the Arab multidimensional poverty index: A theoretical framework

When estimating the multidimensional poverty headcount, MPH, or the Arab multidimensional poverty index, MPI, the analyst assigns a score to each individual in the data set. That score aggregates the data informing 14 indicators (grouped in five dimensions) relevant to the individual's well-being. The MPH or the Arab MPI are then obtained by averaging the scores obtained. Note that many of the indicators relate to negative attributes. Since we aim to model the distribution of each indicator as a distribution of individual attributes in well-being, ill-fare indicators (those that relate to negative attributes) are inverted and interpreted as the absence of the same indicator. A description of the various indicators and their corresponding deprivation cut-offs is provided in table 1.

Table 1. Indicators and scores of the revised Arab multidimensional poverty index

Indicator	Cut-off	Weight
X_1: absence of child mortality for the household	$z_1 = 0$	$d_1 = 1/12$
X_2: the minimum children z-score (wasting and stunting combined) in the household	$z_2 = -2$	$d_2 = 1/12$
X_3: absence of child pregnancy in the household	$z_3 = 0$	$d_3 = 1/12$
X_4: absence of school attendance issues for school-age children in the household	$z_4 = 0$	$d_4 = 1/12$
X_5: absence of schooling gaps for school-age children in the household	$z_5 = 0$	$d_5 = 1/12$
X_6: maximum educational attainment in the household	$z_6 = 11$	$d_6 = 1/12$
X_7: housing is not overcrowded	$z_7 = 0$	$d_7 = 1/12$
X_8: appropriate housing	$z_8 = 0$	$d_8 = 1/12$
X_9: access to improved drinking water	$z_9 = 0$	$d_9 = 1/18$
X_{10}: access to improved sanitation	$z_{10} = 0$	$d_{10} = 1/18$
X_{11}: access to electricity	$z_{11} = 0$	$d_{11} = 1/18$
X_{12}: access to communication assets	$z_{12} = 0$	$d_{12} = 1/18$
X_{13}: access to mobility assets	$z_{13} = 0$	$d_{13} = 1/18$
X_{14}: access to livelihood assets	$z_{14} = 0$	$d_{14} = 1/18$

Source: United Nations, Economic and Social Council for Western Asia, 2020.

The computation of the index is performed in two steps. In the first step, a deprivation cut-off z_k is chosen for each indicator $k \in \{1,2,\dots,14\}$. Then, for each individual $i \in \{1,2,\dots,n\}$, a score is assigned based on the indicators $\mathrm{x}_i = (x_{i1}, x_{i2}, \dots, x_{i15})$ associated with that individual. That score s_i is given by:

(1) $$s_i = \sum_{k=1}^{14} 1\,(x_{ik} \leq z_k) \times d_k.$$

where d_k represents the weight assigned to indicator k, $\sum_{k=1}^{14} d_k = 1$ and $1(\cdot)$ is an indicator function that takes a value of 1 if the argument is true and a value of 0 if it is false.

A second step consists of identifying multidimensional poverty by choosing a cut-off $c \in (0,1)$. That cut-off is fixed at 0.20 for the Arab multidimensional poverty index. An individual is considered poor if their score is greater than or equal to that cut-off. The multidimensional poverty headcount is then given by:

(2) $$MPH = \frac{1}{n}\sum_{i=1}^{n} 1\,(s_i \geq 0.20).$$

The Arab multidimensional poverty index[1] is:

(3) $$MPI = \frac{1}{n}\sum_{i=1}^{n} 1\,(s_i \geq 0.20) \times s_i.$$

It is straightforward to perform the above exercise for each year for which a survey is available. However, surveys are not available for all years. In that context, the analyst must model the values of the two indices for the years without surveys. Below, we propose a modelling approach that provides sufficient flexibility to accommodate a range of assumptions for each indicator.

Alkire and others (2020) propose one approach for forecasting and nowcasting the indices. In the absence of significant shocks, they assume that both indices decrease in accordance with a time-logistic function. In the presence of a major shock, such as the COVID-19 pandemic, they first make a number of assumptions regarding the changes to certain indicators in the last survey year and make corrections to relevant indices. Using those new values, they assume that the indices continue to follow the same logistic trajectory.

In this section, we build on the methodology developed in Alkire and others (2020). Specifically, their use of the most recent data available to take into account major shocks. They also implicitly assume that the correlation among indicators in the most recent year's survey remains stable over time. We formalize that assumption in order to build our own model. One can think of the MPH and the Arab MPI at year t as two functionals of the multidimensional cumulative distribution of indicators, $H_t(\mathrm{x})$ for that year. This multidimensional cumulative distribution function is formally defined as:

1 The individual MPI scores in the associated computer code is equal to $1(s_i \geq 0.20) \times s_i$.

$$(4) \qquad H_t(\mathrm{x}) = H_t(x_1, x_2, \dots, x_{14}) = \Pr[X_1 \leq x_1, X_2 \leq x_2, \dots, X_{14} \leq x_{14} | \text{year } = t].$$

Proposing a model of $H_t(\mathrm{x})$ would allow us to nowcast or forecast the two indices $MPH_t = MPH(H_t(\mathrm{x}))$ and $MPI_t = MPI(H_t(\mathrm{x}))$.

In order to estimate counterfactuals of the multidimensional cumulative distributions $H_t^G(\mathrm{x})$ for the years in which no survey data is available, we draw on the analysis developed in Khaled and others (2020) and build on Sklar's Theorem,[2] which stipulates that $H_t^G(\mathrm{x})$ can be expressed as:

$$(5) \qquad H_t^G(\mathrm{x}) = C_t^G\left(F_{X_1,t}(x_1), \dots, F_{X_K,t}(x_K)\right).$$

i.e. in terms of a K-dimensional copula $C(\cdot)$ and the univariate marginal distribution functions $F_{X_k,t}(x_k)$, $k \in \{1,2,\dots,K\}$. One can generally also recover $H_t^G(\mathrm{x})$ from the copula $C_t^G(\cdot)$ and the marginal distribution functions $F_{X_k,t}(x_k)$, $k \in \{1,2,\dots,K\}$. The main analytical advantage of analysing the problem using a copula approach is that it allows for detailed modelling of each of the indicators $k \in \{1,2,\dots,K\}$ separately and the reconstruction of a counterfactual multidimensional cumulative distribution $H_t^G(\mathrm{x})$ from the counterfactuals of each indicator. The only assumption that one needs to make in this context is that the copula remains stable.[3]

In our context, since we have a number of discrete indicators, the copula function is not unique. Some approaches discussed in the statistical literature provide for the analysis of situations in which certain dimensions are discrete and other are continuous. In order to account for survey weights, we adapt the checkerboard copula estimator developed in Genest, Nešlehová, and Rémillard (2017)[4] using a Hájek weighted estimator (see the annex for an explanation of the estimation procedure). The construction of the model implies estimating the cumulative distribution function for each indicator and the copula for the reference year. We can then use the model to build counterfactuals of the joint distribution and indices when one or more indicators change following an estimated pattern or an assumed exogenous shock.

This approach allows us to decompose the change between two observational years as induced by a change in marginal distributions and a change in the copula. For example, let $\mathrm{H}_{\mathrm{t}_1:\mathrm{t}_0}^{\mathrm{G}}(\mathrm{x})$ represent the multidimensional cumulative distribution function as the copula function of year t_0, $\mathrm{C}_{\mathrm{t}_0}^{\mathrm{G}}(\cdot)$, and the marginal distribution of year t_1, $\mathrm{F}_{\mathrm{X}_\mathrm{k},\mathrm{t}_1}^{\mathrm{G}}(\mathrm{x}_\mathrm{k})$, $\mathrm{k} \in \{1,2,\dots,\mathrm{K}\}$. Using this notation, a change in the Arab MPI for group G between 2011 and 2018 can be rewritten as:

$$(6) \qquad \begin{aligned} MPI(H_{2018}^G(\mathrm{x})) - MPI(H_{2011}^G(\mathrm{x})) = \quad & MPI(H_{2018:2018}^G(\mathrm{x})) - MPI(H_{2011:2018}^G(\mathrm{x})) \\ & + MPI(H_{2011:2018}^G(\mathrm{x})) - MPI(H_{2011:2011}^G(\mathrm{x})) \end{aligned}.$$

[2] For more information on Sklar's Theorem, see Hofert and others (2018).

[3] In Section IV we assess this assumption in the context of the results generated by our illustrative model.

[4] Smith and Khaled (2012) propose an alternative solution, namely a Bayesian latent approach, to model discrete dimensions.

where

(7) $MPI(H^{G}_{2018:2018}(\mathrm{x})) - MPI\big(H^{G}_{2011:2018}(\mathrm{x})\big) = \Delta MPI$ assuming copula stability.

and

(8) $MPI(H^{G}_{2011:2018}(\mathrm{x})) - MPI(H^{G}_{2011:2011}(\mathrm{x})) = \Delta MPI$ induced by change in the copula.

Assuming the stability of the copula allows us to estimate the indices associated with equation (7). If we develop models for each univariate marginal distribution, we can estimate the change corresponding to equation (7). We will come back to this point in our empirical illustration where we will show that the change in the Arab MPI (and most of the change in the MPH) between 2011 and 2018 is captured by the first term of the expression.

An analysis of the 14 indicators of Arab multidimensional poverty reveals that some indicators may affect certain households but may have no impact on others. For example, only households with children under 5 years of age and school-age children may be affected by certain indicators. In order to model the MPH and the Arab MPI effectively, the analyst must decompose the two indices as functions of multiple multidimensional cumulative distributions in order to account for the non-existence of certain indicators for some household types. In our model, we assume that the distribution of household type is the same as it was in 2018. We split the population of n_{2018} households into the following four broad categories:

Group A:

- **Description:** Individuals living in households with children under 5 years of age and school-age children;
- **Number of indicators:** 14;
- **Counterfactual multidimensional cumulative distribution at year** t: $H^{A}_{t:2018}(\mathrm{X})$;
- **Number of observations:** n_{A2018}.

Group B:

- **Description:** Individuals living in households with only children under 5 years of age;
- **Number of indicators:** 12;
- **Counterfactual multidimensional cumulative distribution at year** t: $H^{B}_{t:2018}(\mathrm{X})$;
- **Number of observations:** n_{B2018}.

Group C:

- **Description:** Individuals living in households with only school-age children;
- **Number of indicators:** 13;
- **Counterfactual multidimensional cumulative distribution at year** t: $H^{C}_{t:2018}(\mathrm{X})$;
- **Number of observations:** n_{C2018}.

Group D:

- **Description:** Individuals living in households without children;
- **Number of indicators:** 11;
- **Counterfactual multidimensional cumulative distribution at year t:** $H^D_{t:2018}(\mathrm{X})$;
- **Number of observations:** n_{D2018}.

The model indices for year t are then given by:

$$MPH(H_{t:2018}(\mathrm{X})) = \frac{n_{A2018}}{n_{2018}} MPH(H^A_{t:2018}(\mathrm{X})) + \frac{n_{B2018}}{n_{2018}} MPH(H^B_{t:2018}(\mathrm{X})) + \frac{n_{C2018}}{n_{2018}} MPH(H^C_{t:2018}(\mathrm{X})) + \frac{n_{D2018}}{n_{2018}} MPH(H^D_{t:2018}(\mathrm{X})). \tag{9}$$

and

$$MPI(H_{t:2018}(\mathrm{X})) = \frac{n_{A2018}}{n_{2018}} MPI(H^A_{t:2018}(\mathrm{X})) + \frac{n_{B2018}}{n_{2018}} MPI(H^B_{t:2018}(\mathrm{X})) + \frac{n_{C2018}}{n_{2018}} MPI(H^C_{t:2018}(\mathrm{X})) + \frac{n_{D2018}}{n_{2018}} MPI(H^D_{t:2018}(\mathrm{X})). \tag{10}$$

Let $\mathcal{D}$ indicate the K_G-dimensional domain of integration ($K_G \in \{11,12,13,14\}$). The terms on the right hand side of equations (9) and (10) can be evaluated using the expressions:

$$MPH\left(H^G_{t:2018}(\mathrm{x})\right) = \int_{\mathcal{D}} 1\left(\sum_{k=1}^{K_G} d_k \, 1(x_k \leq z_k) \geq 0.20\right) \mathrm{d}H^G_{t:2018}(\mathrm{x}). \tag{11}$$

and

$$MPI(H^G_{t:2018}(\mathrm{x})) = \int_{\mathcal{D}} \left[1\left(\sum_{k=1}^{K_G} d_k \, 1(x_k \leq z_k) \geq 0.20\right) \times \left(\sum_{k=1}^{K_G} d_k \, 1(x_k \leq z_k)\right)\right] \mathrm{d}H^G_{t:2018}(\mathrm{x}). \tag{12}$$

These multiple integrals can be evaluated by performing a Monte Carlo integration (see the annex for an explanation of the integration procedure).

The model is sufficiently flexible to accommodate counterfactual univariate distribution modelling for different indicators. One can control for covariates of interest using standard econometric methods, including Logit or Probit regressions for dichotomous (0/1) variables, ordered Logit or Probit (or their generalized versions) for ordinal variables, or unconditional quantile regressions (Firpo, Fortin and Lemieux, 2009) or distribution regressions (Chernozhukov, Fernandez-Vál, and Melly, 2013) for continuous variables. One can also follow the approach adopted in Alkire and others (2020) and use simple logistic calibrations. The difference lies in the fact that we do not need to calibrate the trajectory of the MPH or the MPI in our approach. We instead calibrate the change of each univariate distribution at the deprivation cut-off. One can also apply exogenous shocks to model exogenous factors, such as the Covid-19 pandemic or massive destruction of infrastructure following a military operation.

2. A modelling illustration

Since detailed modelling of each indicator is beyond the scope of the present paper, we will use a simple calibration of the univariate distribution of indicators. In more detailed work, we would strongly suggest using detailed econometric models for the indicators that are the focus of the analysis. The analyst should apply simple calibration only on the indicators that are not of immediate interest (with a view to establishing a kind of *ceteris paribus* baseline).

Our illustration uses the 2011 and 2018 UNICEF Multiple Indicator Cluster Surveys for Iraq. Our modelling approach allows us to make a range of assumptions regarding the various indicators. To illustrate the advantage of the approach, we assume that real GDP per capita is the main driver of the proportion of deprived individuals for the following indicators:

1. X_1: absence of child mortality for the household.
2. X_2: the minimum children z-score (wasting and stunting combined) in the household.
3. X_4: absence of school attendance issues for school-age children in the household.
4. X_5: absence of schooling gaps for school-age children in the household.
5. X_7: housing is not overcrowded.

To capture the dynamics in those five indicators from 2011 to 2018, we model deprivation using the following logistic function of real per capita GDP in year t (y_t):

$$(13) \quad F^M_{X_k}(z_k) = \frac{1}{1+e^{-\alpha_k+\beta_k y_t}}.$$

We assume that some indicators change with time and are less subject to per capita GDP fluctuations. We assume that time is the main driver of the proportion of deprived individuals in the following indicators:

1. X_3: absence of child pregnancy in the household.
2. X_6: maximum educational attainment in the household.
3. X_8: appropriate housing.
4. X_9: access to improved drinking water.
5. X_{10}: access to improved sanitation.
6. X_{11}: access to electricity.
7. X_{12}: access to communication assets.
8. X_{13}: access to mobility assets.
9. X_{14}: access to livelihood assets.

To capture the dynamics in those nine indicators from 2011 to 2018, we model deprivation using the following logistic function of time (t):

$$(14) \quad F_{X_k}^{M}(z_k) = \frac{1}{1+e^{-\alpha_k+\beta_k t}}.$$

A vector of 100 parameters must be calibrated in order to capture our simple model of univariate marginal distribution dynamics.[5] Let Ω represent the vector of those 100 calibrated parameters. In line with the approaches adopted in Abdelkhalek and Dufour (1998; 2006) and Dufour, Khalaf and Kichian (2013), we assume that the vector of parameters to be calibrated is a function of several stochastic variables. In other words, $\Omega = \omega(\mathrm{X}_{2011}, \mathrm{X}_{2018})$, a function of the fourteen stochastic indicators in 2011 and 2018. To account for sampling variability, we recalibrate that vector at each bootstrap iteration.

In order to obtain the values of indices for a year t other than 2018, we start by calculating, for each indicator k and each group G, the incidence of deprivation in the relevant indicators for that group, $F_{X_k:t}^{G}(z_k)$, using equations (13) and (14) and integrating any potential exogenous shocks. Using those values, we can generate the dimensional rank, $u_k^z = F_{X_k:t}^{G}(z_k)$. We then use those rank cut-offs to numerically compute the two indices using the estimated copula for 2018. We repeat the procedure for each bootstrap iteration. Note that the calibration part can be substituted in future work by a more sophisticated econometric model for a single attribute or for multiple attributes, using the calibration only for the indicators that we assume follow a constant trajectory.

[5] Vectors of 14×2 for Group A, 12×2 for Group B, 13×2 for Group C and 11×2 for Group D.

3. Our simulation results

Using the 2011 and 2018 Multiple Indicator Cluster Surveys, we estimate the change in the MPH and the Arab MPI. Table 2 illustrates the preliminary estimates of the two indices and their 95 per cent confidence intervals. The preliminary estimates reveal that the incidence of multidimensional poverty decreased by 21 per cent (or by 13 percentage points) between those two years, while the Arab multidimensional poverty index decreased by 27.5 per cent (or by 0.0588).

Table 2. Multidimensional poverty headcount (MPH) and multidimensional poverty index (MPI) in Iraq in 2011 and 2018

	MPH		MPI	
	2011	2018	2011	2018
Estimated index	0.6259	0.4943	0.2142	0.1554
95% confidence interval	[0.6228,0.6290]	[0.4859,0.5027]	[0.2130,0.2153]	[0.1526,0.1582]

Source: Author's preliminary calculations on the basis of data contained in the 2011 and 2018 UNICEF Multiple Indicator Cluster Surveys for Iraq.

To calibrate equations (13) and (14) of the model presented in Section III, we use real per capita GDP and GDP forecasts for 2020, as illustrated in table 3 below. For 2020, the COVID-19 pandemic will affect certain variables via the per capita GDP channel. We also assume that the pandemic will also affect two other indicators. Since a common policy response to the COVID-19 pandemic across countries has been to close schools, we assume an additional 10 per cent increase in attendance issues in addition to the increase caused by the reduction of per capita GDP. Furthermore, since food supply chains may be affected, we also assume an 8 per cent increase in the incidence of child malnutrition (essentially through wasting) in addition to the increase caused by the decrease in per capita GDP.

Table 3. Change in real per capita GDP in Iraq, 2011-2020

	2011	2012	2013	2014	2015
Relative real per capita GDP (2011 = 100)	100	109.78	113.60	110.23	109.27
Real per capita GDP growth rate (per cent)		9.78	3.49	-2.97	-0.86
	2016	2017	2018	2019	2020
Relative real per capita GDP (2011 = 100)	122.33	116.29	112.98	115.32	110.34
Real per capita GDP growth rate (per cent)	11.94	-4.94	-2.84	2.07	-4.31

Source: World Bank, World Development Indicators. 2020 forecast on the basis of data provided by Trading Economics.

Figures 1 to 14 provide preliminary estimates of the change in incidence in Iraq for each of the 14 indicators. The indicators that improved the most between 2011 and 2020 include educational attainment, housing conditions, access to electricity (there are 0 individuals with no access electricity in the 2018 data set), and ownership of communication and livelihood assets. Access to improved sanitation and early pregnancy deteriorated over the same period.

Figure 1. Child mortality

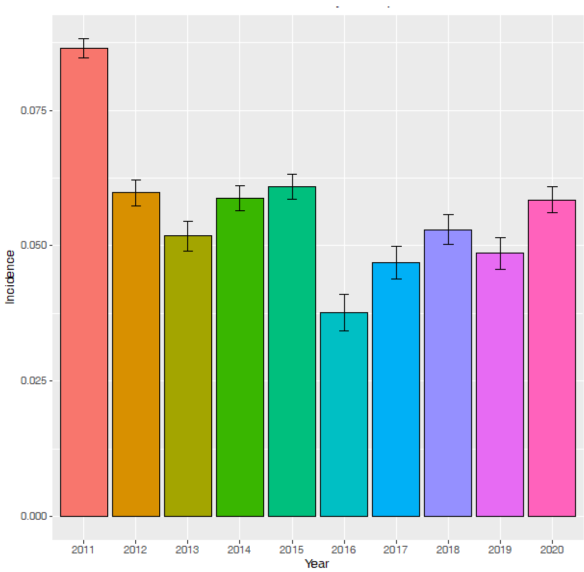


Figure 2. Child malnutrition

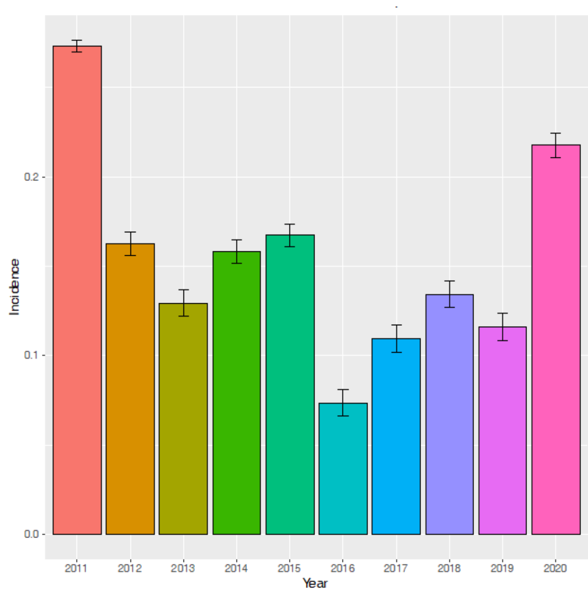


Figure 3. Early pregnancy

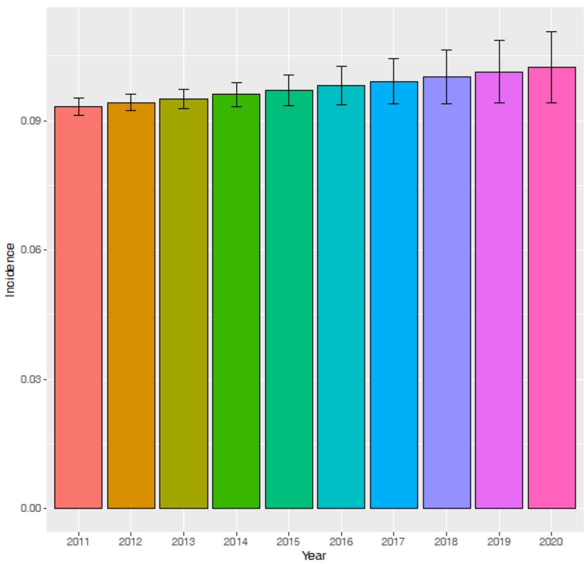


Figure 4. Children not attending school

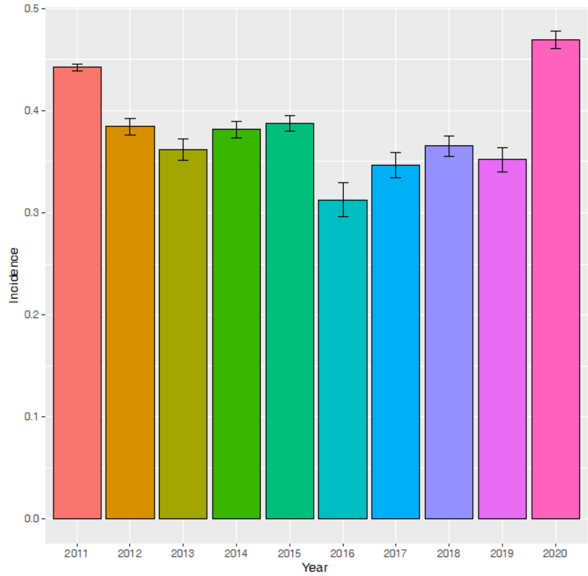


Figure 5. Children with a schooling gap

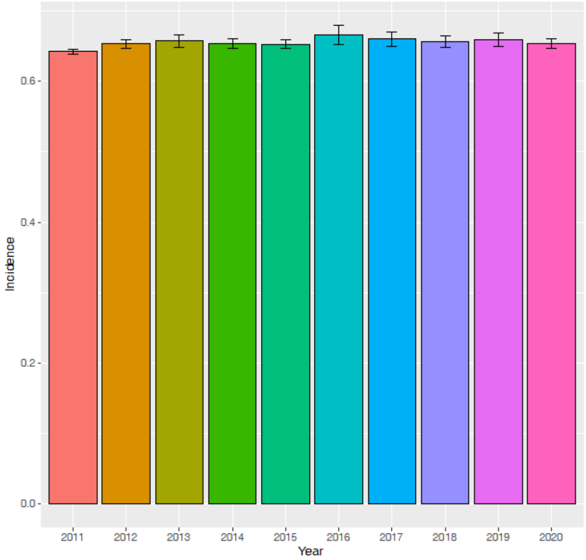


Figure 6. Households with no member having completed secondary education

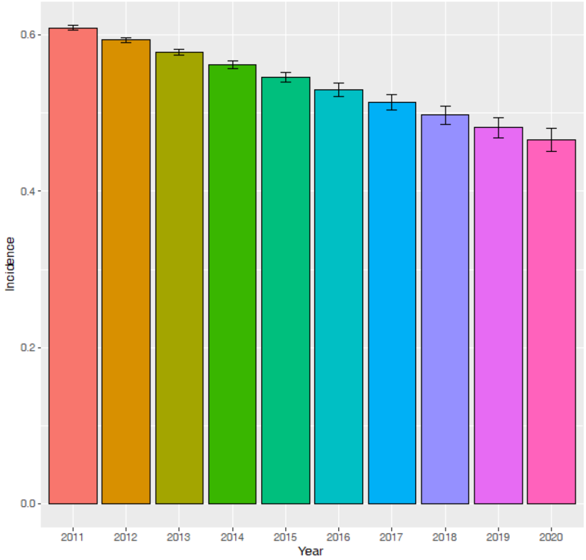


Figure 7. Overcrowded housing

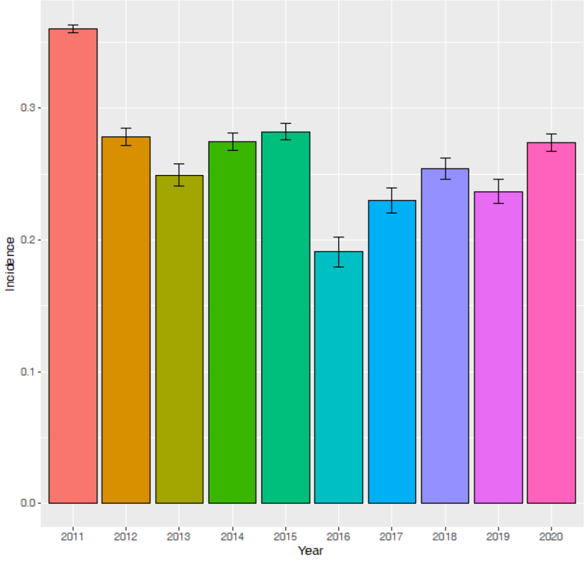


Figure 8. Unsuitable housing

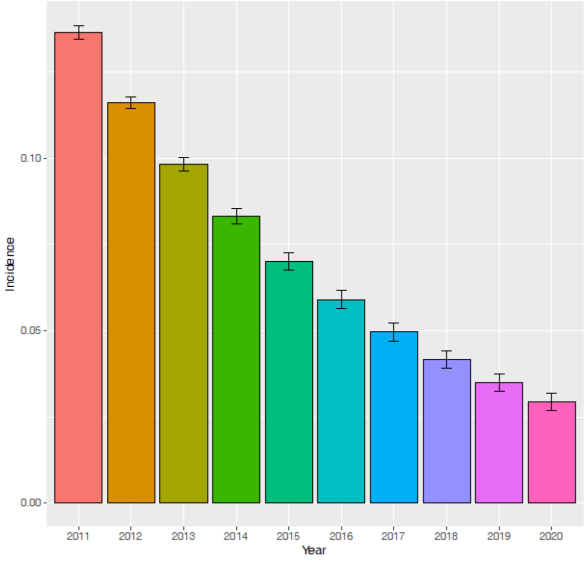


Figure 9. No access to improved drinking water

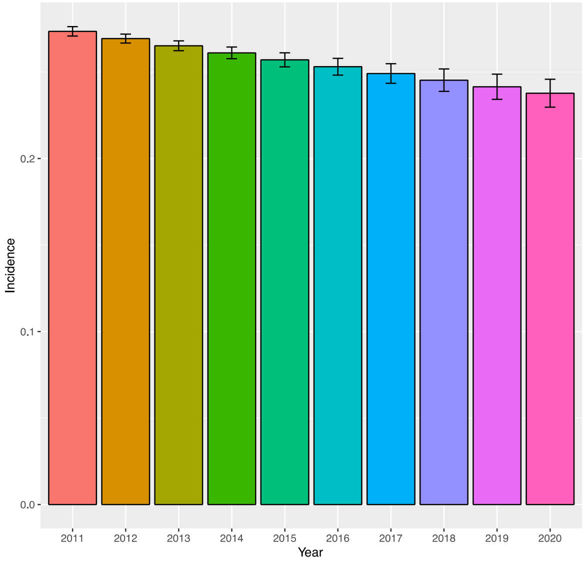


Figure 10. No access to improved sanitation

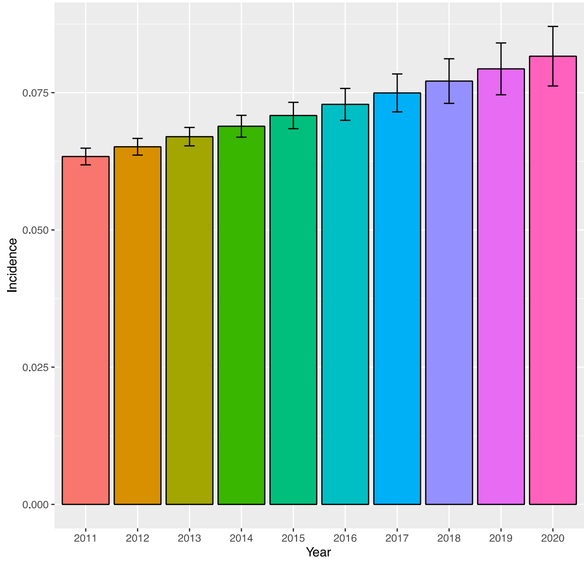


Figure 11. No access to electricity

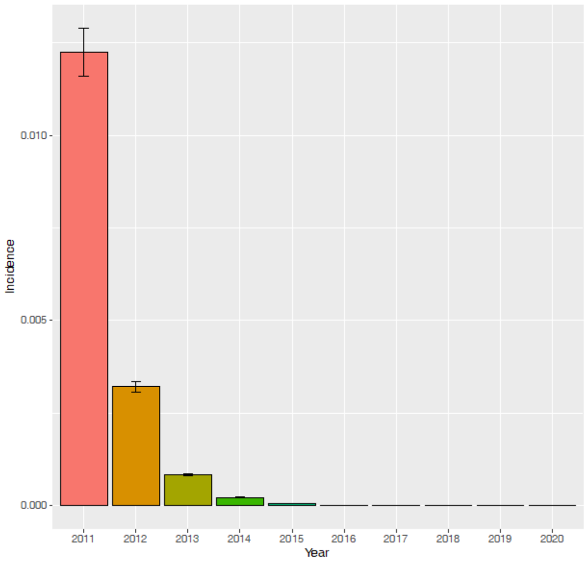


Figure 12. No access to communication assets

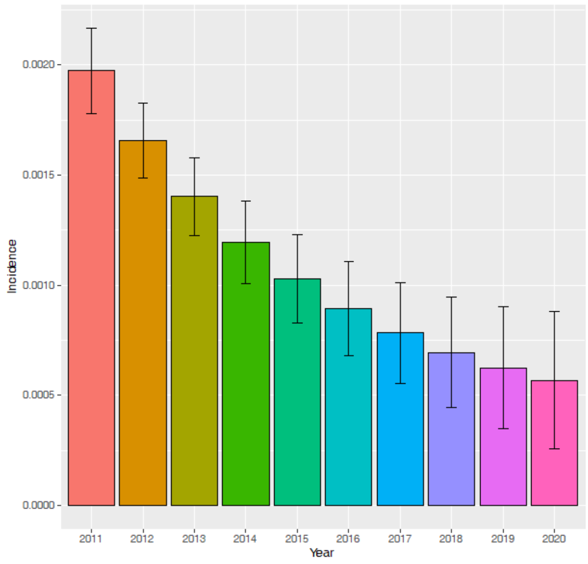


Figure 13. No access to mobility assets (such as a bicycle)

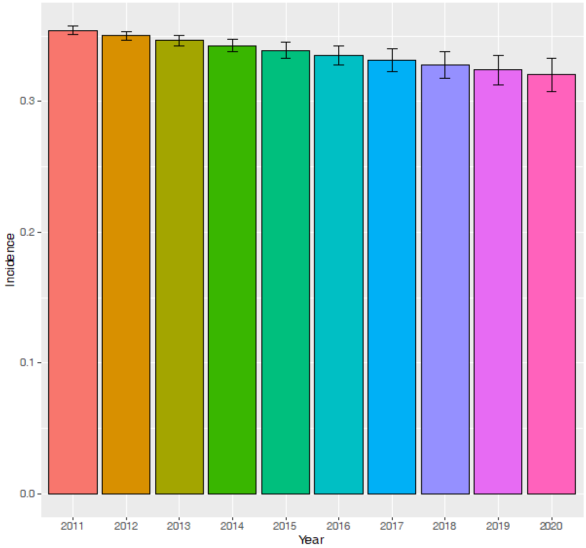
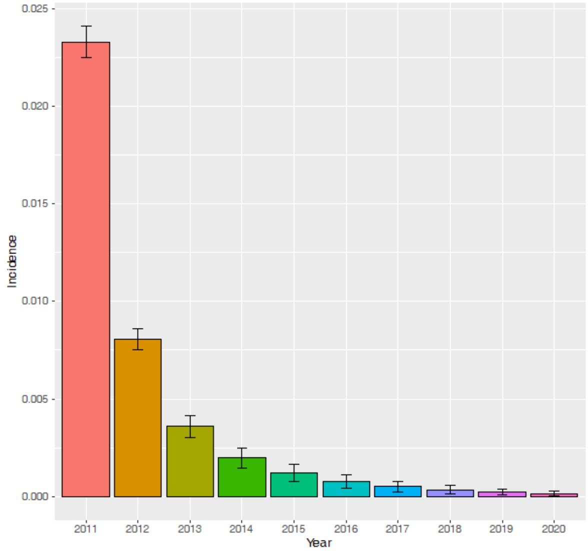


Figure 14. No access to livelihood assets



Figures 1 to 14 should be interpreted with caution, however, particularly as the scale on the y-axis is not the same for all indicators. Figure 15 illustrates the incidence of deprivation for each of the 14 indicators in 2018. As can be seen from that figure, the incidence of deprivation was particularly high in terms of the number of children with a schooling gap and the number of households with no member having completed secondary education, with more than 50 per cent of the population affected. Deprivation in terms of school attendance and mobility asset ownership (such as a bicycle) also affected more than 30 per cent of individuals, while overcrowded housing and no access to improved drinking water affected more than 20 per cent of individuals. Child malnutrition affected almost one child in five.

Figure 15. Deprivation by indicator in 2018

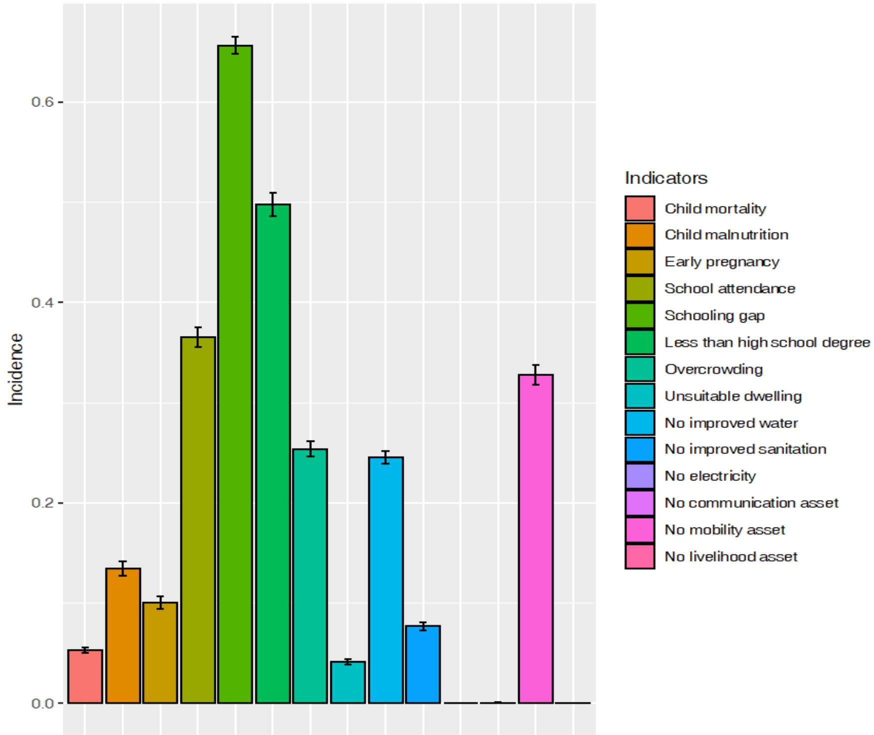


Figure 16. Projected deprivation by indicator in 2020

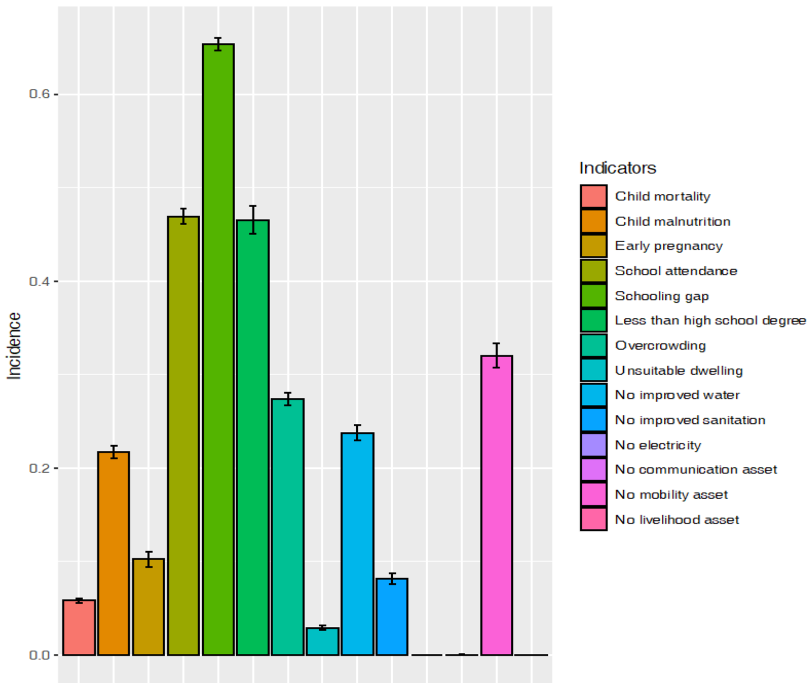


Figure 16 illustrates the incidence of deprivation in 2020 as preliminarily estimated using our model on the basis of our assumptions for each of the 14 indicators. Many of the indicators are affected by the reduction in per capita GDP and other shocks in 2020. Child malnutrition and school attendance are affected disproportionately by the multiple simultaneous shocks.

The model can now be used to aggregate the dynamics of the changes in each indicator in order to generate a dynamic overview of the MPH and the Arab MPI. We estimate the copula for 2018 and introduce the counterfactual distribution of the indicators for each year. Figures 17 and 18 compare the MPH and the Arab MPI values predicted by the model with the values estimated on the basis of Multiple Indicator Cluster Survey datasets. The two indices' numerical values, and their 95 per cent confidence intervals are provided in table 4. For 2018, the values obtained using the model are $MPH(H^G_{2018:2018}(\mathrm{x})) = 0.5278$ and $MPI(H^G_{2018:2018}(\mathrm{x})) = 0.1510$. Those results compare well with the values that can be estimated directly from the datasets. As provided in table 2, $MPH(H^G_{2018}(\mathrm{x})) = 0.4943$ and $MPI(H^G_{2018}(\mathrm{x})) = 0.1554$. The values are therefore relatively close, and the 95 per cent confidence intervals overlap for both indices.

Figure 17. Comparison of the estimated and modelled values of the $\boldsymbol{MPH}$ and Arab $\boldsymbol{MPI}$, 2011

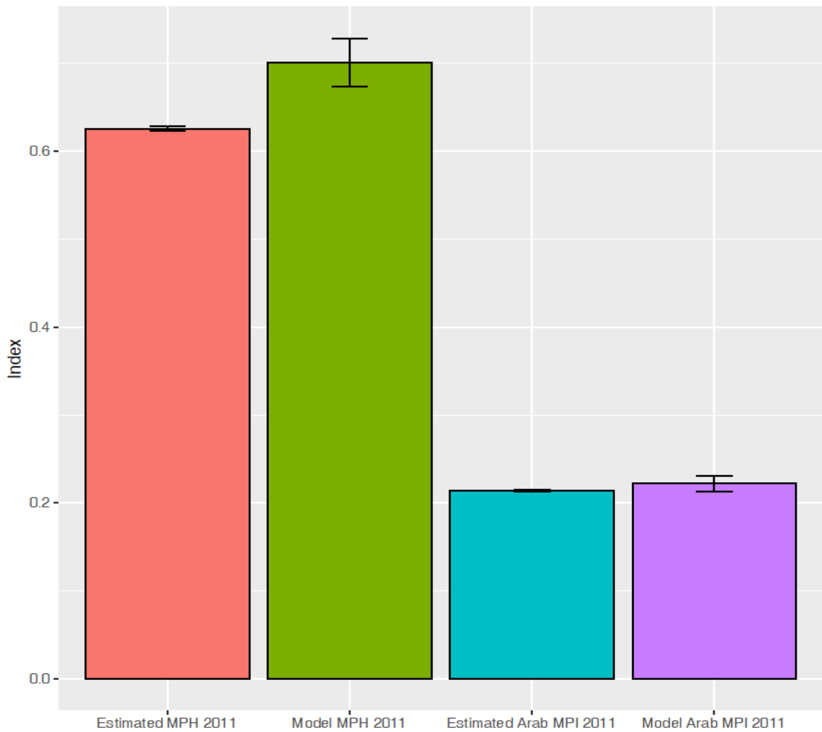
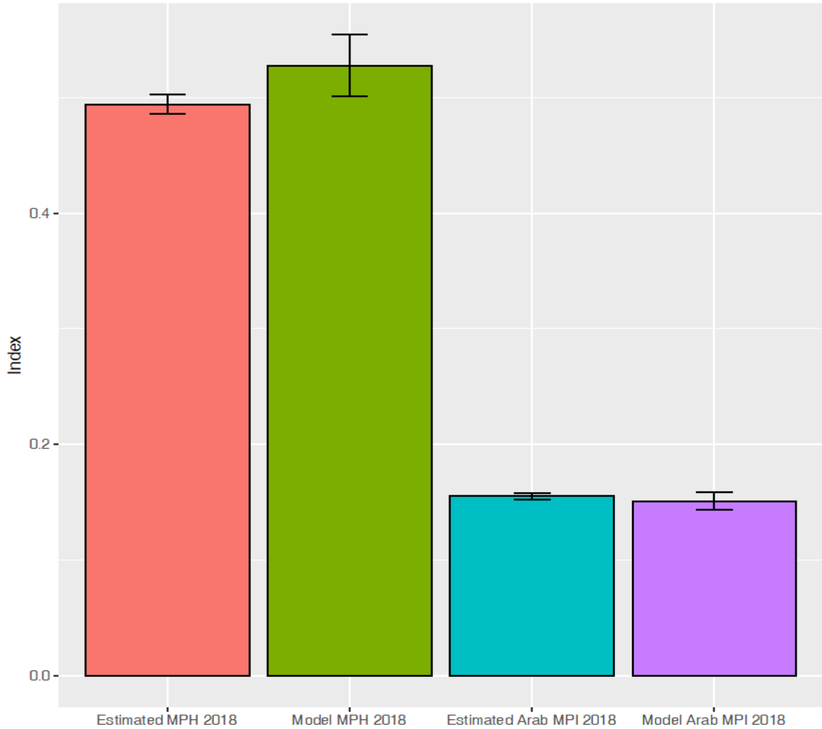


Figure 18. Comparison of the estimated and modelled values of the $\boldsymbol{MPH}$ and Arab $\boldsymbol{MPI}$, 2018



Let us now look at the preliminary estimates for 2011. In this case, keep in mind that we use the marginal distributions of 2011 predicted by the calibrated model. These are, by mathematical construction, the estimated values of 2011. However, the value of the indices predicted by the model for 2011, $MPI(H^G_{2011:2018}(\mathrm{x}))$ and $MPI(H^G_{2011:2018}(\mathrm{x}))$, are obtained by using the estimated copula of 2018. Comparing these values with the indices' estimated values allows us to see how reasonable the assumption of copula stability is over time. For the MPI, the results are striking. Even for the year that is furthest away from 2018 in our illustration, the model predicts a value of $MPI(H^G_{2011:2018}(\mathrm{x})) = 0.2219$ and the index estimated directly from the 2011 data set yields a value of $MPI(H^G_{2011}(\mathrm{x})) = 0.2142$. Since the confidence intervals overlap, we can safely assume that the values of the MPI produced by the model, assuming copula stability over time, is more than reasonable.

Let us now turn to the values of the MPH produced by the model for 2011. The preliminary results are still good but slightly different. The value predicted by the model is $MPH(H^G_{2011:2018}(\mathbf{x})) = 0.7008$ compared to $MPH(H^G_{2011}(\mathbf{x})) = 0.6259$ for the estimated indices. The values are different, but the change is in the same direction. The change in the index is still 76 per cent of the size of the change predicted by the model.

The model performs relatively well in predicting the values at a distance of seven years. It is spot-on for the MPI and produces good values for the MPH. Why would the model perform better for the MPI than for the MPH? Equation (6) reveals that there may be a difference in copulas between the two years. If this difference is more pronounced near deprivation thresholds, this would have a more

significant impact on the MPH than on the MPI because of the discontinuity of the former at the deprivation thresholds.

Table 4. Preliminary model values of multidimensional poverty indices for Iraq, 2011-2020

	2011	2012	2013	2014	2015
$MPH(H_{t:2018}(\mathrm{x}))$	0.7008	0.6137	0.5737	0.5887	0.5817
95% confidence interval	[0.6736,0.7280]	[0.5844,0.6430]	[0.5464,0.6010]	[0.5601,0.6174]	[0.5533,0.6100]
$MPI(H_{t:2018}(\mathrm{x}))$	0.2219	0.1840	0.1674	0.1738	0.1727
95% confidence interval	[0.2134,0.2305]	[0.1752,0.1929]	[0.1594,0.1755]	[0.1651,0.1826]	[0.1643,0.1811]
	2016	2017	2018	2019	2020
$MPH(H_{t:2018}(\mathrm{x}))$	0.4846	0.5112	0.5278	0.4953	0.5869
95% confidence interval	[0.4539,0.5154]	[0.4825,0.5400]	[0.5011,0.5545]	[0.4653,0.5254]	[0.5569,0.6170]
$MPI(H_{t:2018}(\mathrm{x}))$	0.1355	0.1456	0.1510	0.1413	0.1734
95% confidence interval	[0.1265,0.1445]	[0.1371,0.1541]	[0.1432,0.1588]	[0.1326,0.1499]	[0.1646,0.1822]

Source: Author's preliminary calculations on the basis of data contained in the 2011 and 2018 UNICEF Multiple Indicator Cluster Surveys for Iraq.

Figure 19. Projected multidimensional poverty headcount (MPH) for Iraq, 2011-2020

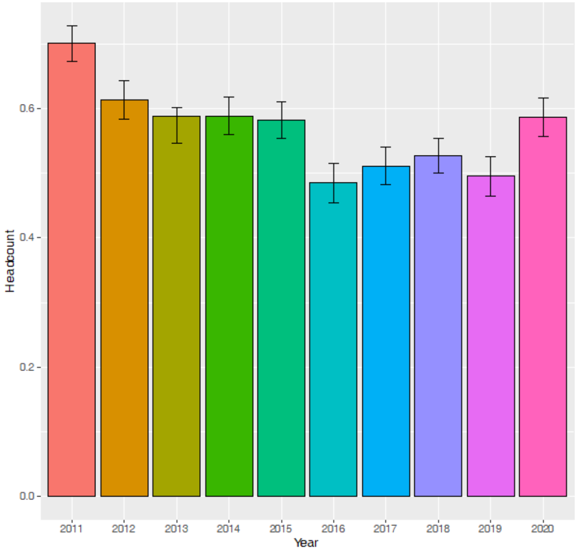
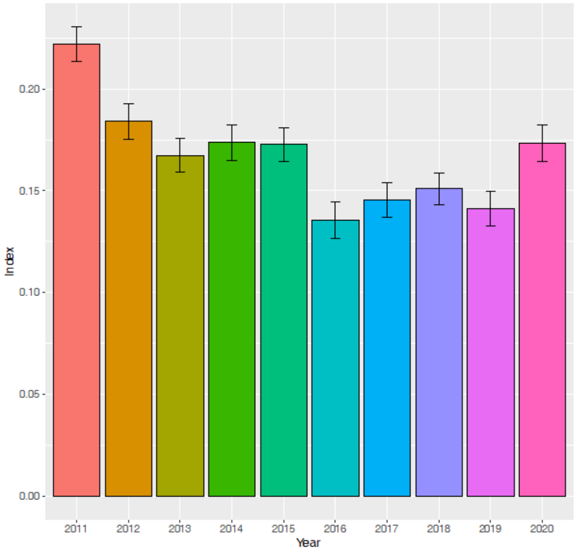


Figure 20. Projected Arab multidimensional poverty index (MPI) for Iraq, 2011-2020



The preliminary MPH and MPI predictions for Iraq by the model over the ten year period is also shown in table 4. Figures XIX and XX also illustrate the results obtained using the model. One result of interest is that, given our assumptions and if one uses the Arab multidimensional poverty index to assess development, the shock associated with the COVID-19 pandemic is equivalent to reversing seven years of development.

It should be remembered that, the longer the time difference between the reference period and the year for which we estimate the copula to model the index, the smaller the proportion of the total variation captured under the stability of the copula assumption of equation (6). The closer we are to the reference year, the larger the proportion of the total change captured under that term. As good values were obtained in 2011 with a seven-year lag, we can safely say that the estimated value is very reasonable under our assumptions regarding the trajectory of each indicator. Conducting more regular surveys in the region would allow researchers to change the reference year for the copula more frequently and produce indices closer to their real values.

4. Conclusion

This paper sets forth an approach for forecasting and nowcasting the MPH and the Arab MPI. In this modelling approach, we assume that the copula remains stable. This is a common assumption in econometrics. The closer we are to a reference year, the more reliable this assumption is likely to be. Our illustration shows clearly that the model can adequately predict variations in the Arab MPI, even with a lag of seven years, in fragile States such as Iraq.

The modelling approach is flexible enough to accommodate different approaches to the estimation (or computation) of the counterfactual univariate distributions for each indicator. This flexibility has many useful policy implications, including that, as demonstrated in Klassen and Lange (2012), we can analyse a single or many indicators in depth using several covariates and then incorporate these into the model. As policies often focus on a single indicator, this analytical approach provides for more accurate predictions of the potential impact of policies on the incidence of deprivation with regard to the indicator of interest and on overall multidimensional poverty indices.

For example, to model the nutrition indicator more effectively, the analyst should use an econometric model for the distribution of that indicator with covariates of interest. Such an approach should take into consideration the educational level of the mother, variables relevant to women's empowerment if those are available, certain regional dummies, and the father's employment status, income or total expenditure. Such an approach would facilitate the building of a counterfactuals distribution of nutritional z-scores. Reweighting methods, such as those proposed in DiNardo, Fortin, and Lemieux (1996), can be used to estimate the counterfactual distribution of a nutrition indicator. Another approach to the same exercise would be the regression distribution approach adopted in Chernozhukov, Fernández-Val and Melly (2013).

In order to model complex shocks, such as the shock stemming from the COVID-19 pandemic, as accurately as possible, use should be made of multidisciplinary teams. For example, using Susceptible, Infectious, or Recovered (SIR) models to capture the dynamics of infection rates and the potential impact on health variables requires expert advice from epidemiologists. Regional covariates must be taken into account in such circumstances. Once modeled correctly, the counterfactual distribution of health variables such as z-scores, can be integrated into the multidimensional poverty index model.

Annex

Estimating the copulas

In order to present the estimator, (see Tillé, 2020), a number of concepts must be introduced. Let ω_i represent the survey weight of individual i. A rank $\hat{u}(x_k)$ associated to a value x_k in indicator k is given by the Hájek estimator of the cumulative distribution:

$$(15)\quad \hat{u}(x_k) = \hat{F}_{X_k}(x_k) = \frac{1}{\sum_{i=1}^{n}\omega_i}\sum_{i=1}^{n}\omega_i\, 1(x_{ik} \leq x_k).$$

Two concepts associated with the above estimator are useful in the estimation of the copula. First, the estimator of the inverse of the cumulative distribution function (CDF) is:

$$(16)\quad \hat{F}_{X_k}^{-1}(u_k) = \inf\{u_k\colon \hat{F}_{X_k}(x_k) \geq u_k\}.$$

The second concept is the left limit of the CDF. It is estimated by:

$$(17)\quad \hat{F}_{X_k}(x_k-) = \frac{1}{\sum_{i=1}^{n}\omega_i}\sum_{i=1}^{n}\omega_i\, 1(x_{ik} < x_k).$$

And the estimator is given by:

$$(18)\quad \hat{C}(\mathrm{u}) = \frac{1}{\sum_{i=1}^{n}\omega_i}\sum_{i=1}^{n}\omega_i \prod_{k=1}^{K}\left\{\lambda_{\hat{F}_k}(u_k)1(u_{ik} \leq u_k) + \left[1 - \lambda_{\hat{F}_k}(u_k)\right]1(u_{ik} < u_k)\right\}.$$

where

$$(19)\quad \lambda_{\hat{F}_k}(u_k) = \begin{cases} \frac{u_k - \hat{F}\left(\hat{F}^{-1}(u_k)-\right)}{\hat{F}(\hat{F}^{-1}(u_k)) - \hat{F}(\hat{F}^{-1}(u_k)-)} & \text{if } \hat{F}\left(\hat{F}^{-1}(u_k)\right) - \hat{F}\left(\hat{F}^{-1}(u_k)\,-\right) > 0. \\ 1 & \text{otherwise.} \end{cases}$$

In the second section of the annex, we also refer to the multidimensional survival function, $\overline{H}(\mathrm{x})$. The copula of this multidimensional survival function, $\overline{C}(\mathrm{u})$, can be estimated using equation (18) but by replacing equations (15), (16), and (17) by their survival counterpart, where the survival is $S_{X_k}(x_k) = 1 - F_{X_k}(x_k)$ for the univariate case (in general this is untrue for the multidimensional case, and hence the copula of the survival should be estimated).

Numerical integration of the indices

In order to numerically compute the indices from the counterfactual distributions, we simulate 1,000 K-dimensional vectors, $\mathrm{u_i}$ $(i \in \{1,2,\ldots,1000\})$, for each household type. To each one of these vectors, we associate $\hat{\psi}(\mathrm{u_i}) = \widehat{\mathrm{Pr}}[u \in \mathcal{B}^K(\hat{\mathrm{u}}_\mathrm{i}^-, \hat{\mathrm{u}}_i)]$, where $\mathcal{B}^K(\hat{\mathrm{u}}_\mathrm{i}^-, \hat{\mathrm{u}}_i) = \prod_{k=1}^K [\hat{\mathrm{u}}_\mathrm{i}^-, \hat{\mathrm{u}}_i]$, is the K-dimensional hyperbox defined between $\hat{\mathrm{u}}_\mathrm{i}^-$ and $\hat{\mathrm{u}}_i$, and $\hat{\mathrm{u}}_\mathrm{i}^- = \hat{\mathrm{u}}_i - (\varepsilon, \ldots, \varepsilon)$. $\varepsilon > 0$ should be chosen small enough. We use $\varepsilon = 0.0001$ in our numerical application. The probability of an individual being in the hyperbox can be expressed as: $\hat{\psi}(\hat{\mathrm{u}}_i) = \mathrm{Pr}[\hat{\mathrm{u}}_\mathrm{i}^- <_K u \leq_K \hat{\mathrm{u}}_i]$, where $<_K$ and $\leq_K$ are the vectors element-wise operators. From there, it is straightforward that this expression is equivalent to:

(20) $\hat{\psi}(\hat{\mathrm{u}}_i) = \hat{C}(\hat{\mathrm{u}}_i) \times \hat{\bar{\hat{C}}}(\hat{\mathrm{u}}_\mathrm{i}^-)$.

The two terms on the right-hand side of this equation can be estimated using the copula estimation described in the previous section of the annex.

The value of the indices are given by:

(21) $\widehat{MPH}(H_{t:2018}^G(\mathrm{x})) = \frac{1}{\hat{\varphi}} \sum_{i=1}^{1,000} \hat{\psi}\,(\mathrm{u_i}) 1\left(\sum_{k=1}^K d_k\, 1(u_{ik} \leq \hat{F}_{X_k t}^G(z_k)) \geq 0.20\right)$.

and

$\widehat{MPI}\left(H_{t:2018}^G(\mathrm{x})\right) =$

(22) $\frac{1}{\hat{\varphi}} \sum_{i=1}^{1,000} \hat{\psi}\,(\mathrm{u_i}) \left[1\left(\sum_{k=1}^K d_k\, 1\left(u_{ik} \leq \hat{F}_{X_k t}^G(z_k)\right) \geq 0.20\right) \times \sum_{k=1}^K d_k\, 1\left(u_{ik} \leq \hat{F}_{X_k t}^G(z_k)\right)\right]$.

where

$\hat{\varphi} = \sum_{i=1}^{1,000} \hat{\psi}\,(\mathrm{u_i})$.

References

Abdelkhalek, Touhami and Jean-Marie Dufour (1998). Statistical inference for computable general equilibrium models, with application to a model of the Moroccan economy, *Review of Economic and Statistics*, vol. 80(4), pp. 520-534.

Abdelkhalek, Touhami and Jean-Marie Dufour (2006). Confidence regions for calibrated parameters in computable general equilibrium models, *Annales d'Économie et de Statistique*, vol. 81, pp. 1-31.

Alkire, Sabine. and James Foster (2011). Counting and multidimensional poverty measurement, *Journal of Public Economics*, vol. 95, issues 7-8, pp. 476-487.

Alkire, Sabine and others (2020). On track or not? Projecting the global multidimensional poverty index, *Research in Progress Series* 58a, Oxford Poverty & Human Development Initiative (OPHI).

Alkire, Sabine, and Maria Emma Santos (2014). Measuring acute poverty in the developing world: Robustness and scope of the multidimensional poverty index, *World Development*, vol. 59, pp. 251-274.

Atamanov, Aziz and others (2020). Measuring monetary poverty in the Middle East and North Africa (MENA) region: Data gaps and different options to address them, Policy Research Working Paper, No. WPS 9259. Washington, D.C.: World Bank Group.

Chernozhukov, Victor, Iván Fernández-Val and Blaise Melly (2013). Inference on counterfactual distributions, *Econometrica*, vol. 81, No. 6 (November), pp. 2205-2268.

DiNardo, John, Nicole Fortin and Thomas Lemieux (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach, *Econometrica*, vol. 64, No. 5 (September), pp. 1001-1044.

Dufour, Jean-Marie, Lynda Khalaf and Maral Kichian (2013). Identification-robust analysis of DSGE and structural macroeconomic models, *Journal of Monetary Economics*, vol. 60(3), pp. 340-350.

United Nations, Economic and Social Commission for Western Asia (2020). *Proposal for a Revised Multidimensional Poverty Index for Arab Countries*. E/ESCWA/CL2.GPID/2020/TP.11, L20-00300.

Firpo, Sergio, Nicole Fortin and Thomas Lemieux (2009). Unconditional quantile regressions, *Econometrica*, vol. 77, No. 3 (May), pp. 953-973.

Genest, Christian, Johanna Nešlehová, and Bruno Rémillard (2017). Asymptotic behavior of the empirical multilinear copula process under broad conditions, *Journal of Multivariate Analysis*, vol. 159, pp. 82-110.

Hofert, Marius and others (2018). *Elements of Copula Modelling with R*, Springer.

Khaled, Mohamad and others (2020). A simple unidimensional representation of multidimensional inequality, with an application to the Arab region, mimeo (available from the author).

Klasen, Stephan and Simon Lange (2012). Getting progress right: Measuring progress towards the MDGs against historical trends, Courant Research Centre: Poverty, Equity and Growth–Discussion Papers 87.

Ram, Rati (2020). Attainment of multidimensional poverty target of sustainable development goals: a preliminary study, forthcoming in *Applied Economics Letters*, 31 May 2020.

Smith, Michael and Mohamad Khaled (2012). Estimation of copula models with discrete margins via Bayesian data augmentation, *Journal of the American Statistical Association*, vol. 107, issue 497, pp. 290-303.

Tillé, Yves (2020). *Sampling and Estimation from Finite Populations*, Wiley Series in Probability and Statistics.

The present publication is produced under the project on Strengthening Social Protection for Pandemic Responses, jointly implemented by the United Nations Department of Economic and Social Affairs, UNCTAD and the United Nations regional commissions, and funded by the United Nations Development Account. The project aims to build national capacity to design and implement social protection policies, with a gender perspective, for a rapid recovery from the COVID-19 pandemic; and to increase resilience to the negative impact of future exogenous shocks, especially among the most vulnerable populations. The project comprises the following three workstreams: building social protection capacities; advancing care economy; and guiding poverty reduction.

Strengthening
Social Protection for
Pandemic Responses
Building Social Protection Capacities

Strengthening
Social Protection for
Pandemic Responses
Guiding Poverty Reduction

Strengthening
Social Protection for
Pandemic Responses
Advancing Care Economy

21-00224