IARIW 2025

IARIW 2025

Thursday, October 2 & Friday, October 3

Using Cross-survey Imputation to Estimate Food Insecurity Prevalence in Data-Scarce Environments

Jad Chaabana
(Doha Institute for Graduate Studies, Doha, Qatar)

Fatima Zahra Bendriss
(Doha Institute for Graduate Studies, Doha, Qatar)

jad.chaaban@dohainstitute.edu.qa

Paper prepared for the IARIW–World Bank–UEB/VNU Conference on "Improving Well-being Measurement in Data-challenged Environments in Developing Countries for Better Evidence-based Policies" October 2-3, 2025

Session 3B Well-being
Time Slot: Thursday, October 2, 3:30- 5:30 PM

Using cross-survey imputation to estimate food insecurity prevalence in data-scarce environments

Jad Chaaban[a], *Doha Institute for Graduate Studies, Doha, Qatar*

Fatima Zahra Bendriss, *Doha Institute for Graduate Studies, Doha, Qatar*

Abstract

The global fight to reduce hunger requires accurate data collected and analyzed regularly. This is not the case in many countries especially developing and conflict-stricken ones, where food security - specific surveys are not consistently implemented. We use a cross-survey imputation methodology to estimate the prevalence of food insecurity in datasets that do not collect this information, including estimates at the sub-national level. We apply this framework to the case of Lebanon, where we use recent freely available microdata from two nationally representative surveys conducted in the country between 2023 and 2024: FAO's Food Insecurity Experience Scale (FIES) survey (2023), and the Arab Barometer Wave VIII survey (2024). Our results show that imputed food insecurity prevalence indicators can be obtained in a robust manner with few restrictions. We also provide an estimation of sub-national food insecurity prevalence in the country.

Keywords: Food security; Data imputation; FIES; Surveys; Lebanon.

[a] Corresponding author, Doha Institute for Graduate Studies, Doha, Qatar.
PO Box: 200592, Doha, Qatar. Tel: +974 40356914. jad.chaaban@dohainstitute.edu.qa.

1. Introduction

The availability of timely and disaggregated food insecurity measures especially at the sub-national level is of crucial importance in the fight towards reducing hunger. Yet many countries lack consistent and detailed data on food security indicators, especially in contexts of crises and conflicts. The latest available data on the progress towards achieving the hunger and food insecurity target (SDG Target 2.1) shows that the most recent figures for the prevalence of moderate or severe food insecurity are available up to 2022 only. Moreover, these indicators were missing for more than 40 countries and replaced by contemporaneous model estimates (FAO et al., 2024). Several African countries where food insecurity is very high are included in this list, in addition to countries in other regions witnessing conflict or political turmoil (Syria, Sudan, Yemen…). Obtaining and/or updating food insecurity indicators is becoming crucial, especially in data-scare environments and crisis situations where food security and livelihoods deteriorate quickly.

Cross-survey imputation methods have been used to estimate poverty across geographical areas in the context of poverty mapping exercises or trends across time periods in the context of repeated cross-section or panel data (Dang and Verme, 2023, Beltramo et al., 2024). Yet to date these methods have not been used to estimate food insecurity prevalence within developing countries. There is only one recent research that used cross-survey multiple imputation to assess within-state food security patterns in the United States (Bartfeld and Reed-Jones, 2024). This study benefited from an elaborate set of variables at the micro level with surveys sharing the same sampling frame. Our main contribution in this paper is to present and test a framework for cross-survey multiple imputation of food security indicators that is adapted for settings where few variables are collected, and where various surveys could originate from different samples.

The methodology relies on first estimating an empirical model of the determinants of food insecurity at the individual level, using factors like income, gender, age, household size and composition. The estimated parameters are then used in out-of-sample predictions of food insecurity prevalence rates in another dataset that contains similar variables, but where the main indicator of interest is missing. The common variables are standardized using a statistical adjustment to ensure compatibility across both surveys. The empirical estimation also generates prediction errors and confidence intervals which are useful in evaluating the robustness of the findings.

We apply this framework to the case of Lebanon, where we use recent microdata from two nationally representative surveys conducted in the country: FAO's Food Insecurity Experience Scale (FIES) survey, and the Arab Barometer Wave VIII survey. Both surveys were conducted fairly recently and contain several demographic, regional and income variables that are useful to implement the cross-survey imputations. The main dependent welfare variable we use is derived from the FIES. It is a widely used experience-based metric of food insecurity, which relies on people's direct responses to questions about their experiences facing constrained access to food (Saint Ville et al., 2019). It focuses on individual's self-reported food-related behaviors and experiences, and the survey contains several demographic and economic variables that are important determinants of food insecurity.

The rest of the paper is organized as follows: Section 2 provides the overall theoretical and empirical framework for our cross-survey imputation method, and describes the datasets we use from Lebanon. Section 3 presents our main findings and estimates, including national food insecurity estimates in addition to subnational ones. The latter are benchmarked against other relevant recent subnational food insecurity indicators to assess the validity of our imputations. And Section 4 concludes the paper with a discussion of the results and a layout of future related research.

2. Materials and methods

2.1. Theoretical framework for cross-survey imputation

The analytical framework for cross-survey imputation is adapted from the poverty imputation methodology developed by Dang et al. (2017). Assume that we have two surveys, denoted by survey 1 and survey 2, where survey 1 contains food security related score data and survey 2 does not. These two surveys can be either in the same period or in different periods. Let x_j be a vector of common characteristics between the two surveys, where j indicates the survey reference and y_j represents the food security score. The objective is to impute the missing food security score data y_2 in survey 2, given that the food security score data is available in survey 1 *only*, and the survey characteristics x_j are available in *both* surveys.

We assume that the linear projection of the food security score on characteristics (x) for survey 1 is given by a cluster random-effects model:

$$y_1 = \beta_1' x_1 + \nu_1 + \varepsilon_1 \qquad (1)$$

And a similar linear projection of the food security score y_2 in survey 2:

$$y_2 = \beta_2' x_2 + \nu_2 + \varepsilon_2 \qquad (2)$$

where the cluster random effects and the error terms are assumed uncorrelated with each other and to follow a normal distribution $\nu_j \mid x_j \sim N(0, \sigma^2_{\nu_j})$ and $\varepsilon_j \mid x_j \sim N(0, \sigma^2_{\varepsilon_j})$; and βj are the vector of coefficients, for periods $j = 1, 2$.

Let th_2 be the food security threshold in period 2 (below which an individual with a low food security score is classified as food insecure); if y_2 existed the food insecurity rate FI_2 in this period could be estimated with the following:

$$FI(y_2 \leq th_2) \quad (3)$$

where FI(.) is the probability function that gives the percentage of the population that are under the food security threshold th_2 in survey 2.

Dang et al. (2017) present a critical assumption which ensures that the sampled data in survey 1 and survey 2 are each representative of the target population, thus making survey-to-survey imputation possible:

"Assumption 1: Let x_j denote the values of the variables observed in survey j, for j = 1, 2, and let X_j denote the corresponding measurements in the population. Then x_j are consistent measures of X_j for all j (i.e. $x_j = X_j$ for all j)."

This assumption is likely to be met in surveys of the same design and sample frame. In this case, and since the estimated parameters are obtained using a different survey from the target survey, we can use simulation to estimate Equation 2 as follows:

$$\hat{y}_{2,s}^{1} = \hat{\beta}_1{}'x_2 + \tilde{\hat{v}}_{1,s} + \tilde{\hat{\varepsilon}}_{1,s} \quad (4)$$

where $\hat{\beta}_1$,$\hat{\sigma}_{v_1}^2$,$\hat{\sigma}_{\varepsilon_1}^2$ represent the estimated parameters obtained from Equation 1 and $\tilde{\hat{v}}_{1,s}$ and $\tilde{\hat{\varepsilon}}_{1,s}$ represent the sth random draw from their estimated distributions, for s = 1,…, S. And based on the estimated food security score parameters based on survey 1 and the data in survey 2, the food insecurity rate in period 2 can be predicted as:

$$FI(y_2) = FI(y_2^1) \quad (5)$$

More specifically, the imputed food insecurity rate FI_2 and its variance in the target survey are estimated as:

$$\widehat{FI}_2 = \frac{1}{S}\sum_{s=1}^{S} FI\ (\hat{y}_{2,s}^{1} \leq th_1) \quad (6)$$

$$\mathrm{V}(\widehat{\mathrm{FI}}_2) = \frac{1}{\mathrm{S}}\sum_{\mathrm{s=1}}^{\mathrm{S}} \mathrm{V}\ (\widehat{\mathrm{FI}}_{2,\mathrm{s}}|\mathrm{x}_2) + \mathrm{V}(\frac{1}{\mathrm{S}}\sum_{\mathrm{s=1}}^{\mathrm{S}} \widehat{\mathrm{FI}}_{2,\mathrm{s}}|\mathrm{x}_2) \quad (7)$$

(See Dang et al. 2017 for the detailed proofs related to the poverty function imputation).

The intuition behind this food insecurity imputation method is that we predict the food security score variable in the target survey based on the estimated food security score parameters (and the error term) and their distributions using Eq. (1). Once we obtain the predicted distribution of the food security score variable, we use it to estimate the food insecurity rate as in Eq. (6).

For surveys of different designs, Dang et al. (2017) provide a method to 'standardize' the distributions of the variables in survey 2 by those in survey 1 in Proposition 2 in their paper:

"Proposition 2: Standardizing common variables in surveys of different design: Assume that survey 2 has the same design over time, is collected more frequently than survey 1, and that the time periods that data from the former are available include the periods that data from the latter are available. Assume further that the overlapping variables between the two surveys follow a normal distribution such that $x_{1t} \sim N(\mu_{1t}, \sigma_{1t}^2)$ and $x_{2t} \sim N(\mu_{2t}, \sigma_{2t}^2)$, for $t = 1,\ldots,$ T. We can standardize the variables in survey 2 according to survey 1 for both the overlapping periods and other periods as follows:

i) For the overlapping period *t*, the standardized variables $x_{2\to1,t}$ in survey 2 are given by

$$x_{2\to1,t} = (x_{2t} - \mu_{2t})\frac{\sigma_{1t}}{\sigma_{2t}} + \mu_{1t}$$

ii) For period *t'* where only data from survey 2 are available, assuming further that the standardized changes between time *t* and time *t'* are the same for the variables x in survey 1 and survey 2 (that is, $\frac{\sigma_{1t}}{\sigma_{2t}}(\mu_{2t'} - \mu_{2t}) = (\mu_{1t'} - \mu_{1t})$) and the variances of the variables x are the same between different rounds of the same survey (that is $\frac{\sigma_{jt}}{\sigma_{jt'}} = 1$, for j = 1, 2). The standardized variables $x_{2\to1,t'}$ in survey 2 are given by

$$x_{2\to1,t'} = (x_{2t'} - \mu_{2t})\frac{\sigma_{1t}}{\sigma_{2t}} + \mu_{1t} \quad \text{''}$$

(Proof: See Dang et al. 2017).

This standardization procedure ensures that the transformed variables used in the imputation procedure satisfy Assumption 1 above.

2.2. Empirical strategy

In light of the analytical framework detailed above, cross-survey imputation of food insecurity scores involves the following steps:

a) Identify the list of independent variables that can act as good predictors for the food insecurity prevalence indicator, and which are present in both surveys (the original one and the target one).
b) Compare of the distribution of the common variables between the two surveys and test for their differences, by testing if the means of the variables are statistically different between these surveys.
c) If the differences are statistically significant, then the distributions of the variables in survey 2 should be 'standardized' by those in survey 1, as detailed previously.
d) Implement the imputation of the food security raw score and the prevalence of food insecurity, by using a standard imputation framework (such as the Multiple Imputation MI method), and by choosing an appropriate threshold for the various food insecurity prevalence rates.

The list of common predictors of food insecurity includes demographics, socio-economic conditions, environmental variables as well as government policies. Among these, economic factors, especially income, are among the most important (Bartfeld et al. 2024; Mahmood et al. 2023, Jubayer et al. 2023). Food insecurity risks are also shaped by demographic characteristics: The larger the household size and number of dependents, the more likely a household is to be food insecure as the provision of food for numerous members becomes difficult (Mahmood et al., 2023). Gender is also a stratifying factor, as female-headed households are considered more likely to be food insecure (Bartfeld et al., 2024; Mahmood et al., 2023), while education has been found to be an attenuating factor of food insecurity. Age composition also affects food security status: Households that have at least one elderly member tend to be less at risk, perhaps due to steady income sources coming from pensions or other forms of support received (Bartfeld et al., 2024). Moreover, ownership of assets; especially agricultural and property; is considered an important buffer (Mahmood et al., 2023), similar to human capital assets such as the education level of the household head (Bartfeld et al., 2024; Mahmood et al., 2023).

To test for the similarity of the common variables' distribution across both surveys, a mean comparison two sample t-tests can be used (Wooldridge, 2010). If the differences are statistically significant, especially when the two surveys originate from different sampling frames, then the standardization of the common variables should be implemented as detailed in Dang et al. (2017).

The multiple imputation (MI) across-survey method involves standard MI methods (van Buuren, 2007, Raghunathan et al., 2001) which impute multiple estimates of the target dependent variable (food insecurity) in the second survey based on estimating a predictive model for this variable in the first survey, where regressors include common variables across both surveys. The estimates in the second survey are each a random draw from the distribution generated by the model parameters.

2.3. Application to datasets from Lebanon

2.3.1. FAO Food Insecurity Experience Scale (FIES) survey

We use two datasets to carry out the imputation method, with the Food and Agriculture Organization's (FAO) Food Insecurity Experience Scale (FIES) survey playing a central role in identifying food insecurity. Since its introduction in 2014, the FIES has been implemented in over 140 countries, and delivers consistent and reliable data that allows for meaningful comparisons across diverse populations (Ballard, Kepple, Cafiero, 2013).

The FIES survey uses a module of eight carefully questions to assess food insecurity at both the household and individual levels. These questions capture key experiences such as reduced food intake, disrupted eating patterns, poor dietary quality, and hunger. Responses to these questions produce a raw score ranging from 0 to 8, where higher scores correspond to greater severity. We use thresholds for moderate and severe Food Insecurity (FI) following the FAO-VoH thresholds as food secure (raw score = 0–3), moderate FI (raw score = 4–6), and severe FI (raw score = 7–8). The FIES and its thresholds have been recently assessed for Arab countries including Lebanon and have been found to be a valid measure of food insecurity in these countries (Sheikomar et al., 2021).

In Lebanon, the 2023 FIES survey provides extensive national coverage and targets individuals aged 15 and older (FAO, 2024). The sampling frame encompasses both urban and rural populations, ensuring representation of the civilian, non-institutionalized demographic. Data collection, which was implemented between June and July 2023, employed a dual-frame approach, capturing responses from individuals with access to landlines and/or mobile phones. The dataset contains, in addition to the FIES module, demographic variables related to number of adults and children in the household, age, education, area (urban/rural), gender, income and degree of urbanization.

2.3.2. Arab Barometer survey

The second 'target' dataset in this study is the Arab Barometer Wave VIII survey (2023–2024), which is the most recent microdata representative survey in Lebanon. The Arab Barometer conducts in-depth public opinion surveys across the Middle East and North Africa (MENA), creating one of the most extensive and long-standing collections of publicly available data on the views and values of people in the region. The Arab Barometer survey is nationally representative, targeting citizens aged 18 and older from both urban and rural areas, and uses a probability-based sampling approach, by selecting one individual from each household to ensure a balanced and accurate representation. This rich dataset provides important insights into various facets of life in the MENA region, offering valuable information to policymakers, researchers, and civil society organizations. For this study, the Arab Barometer dataset is especially useful because it complements the FAO dataset by providing detailed information on the factors explaining food insecurity at the subnational governorate level—an essential element missing from the FIES survey. This more detailed data enables a better understanding of regional differences in food insecurity, facilitating a more targeted and localized analysis. Data collection was implemented in Lebanon between February and April 2024, on a nationally representative sample of 2403 individuals.

The examination of the variables in both the FAO 2023 and the Arab Barometer 2024 surveys shows that the common variables across these two surveys include: Age, Gender, Education, Number of Adults in Household, Number of children (below 15) in Household, and Income. Since the FAO dataset contains only income quintiles, we convert the income variable in the Arab Barometer dataset into quintiles also.

3. Results

3.1 Comparison of the two surveys

The core assumption of our imputation approach is that key independent variables in our two datasets stem from a common population. To assess the validity of this assumption, it is important to compare the main variables age, gender, basic education (which equals 1 if the respondent has elementary education or less), household size, the share of children in the household, and income (categorized into five quantiles, both in the FAO and Arab Barometer datasets). Overall, Table 1 shows that the means of these variables are found to be similar across the two datasets for only 3 out of the 6 core variables, which implies the need to standardize the variables in survey 2 (Arab Barometer) by those in survey 1 (FAO FIES).

Table 1: Two-sample t test with equal variances for the main variables used

	FAO 2023	AB 2024	Mean1	Mean2	diff	St Err	t value	*p value*
Age	1000	2403	45.055	42.233	2.822**	.585	4.85	*0.00*
Gender	1000	2403	0.496	.495	.001	.019	0	*.985*
Household size	1000	2403	4.090	3.776	.315**	.062	5.05	*0.00*
Share children	1000	2403	0.175	.178	-.004	.008	-.45	*.671*
Having basic education	1000	2403	0.339	.342	-.003	.018	-.15	*.882*
Quintile of Income	1000	2403	3.115	2.913	.202**	.054	3.75	*0.00*

H_0: Mean1=Mean2. **Reject H_0, there is a statistically significant difference in the means.

3.2. Imputation of food insecurity prevalence

The empirical model we use to predict food insecurity in the first survey is specified as follows:

$$RS = \beta_0 + \beta_1 \text{ GENDER} + \beta_2 \text{ AGE} + \beta_3 \; AGE^2 + \beta_4 \text{ EDU} + \beta_5 \text{ HHsize} + \beta_6 \; HHsize^2 + \beta_7 \text{ Sh\_child} + \beta_8 \text{ INC} + u$$

where:

- *RS*: Raw Score of the FIES module, which measures the sum of answers to the 8 FIES questions. RS ranges from 0 to 8.
- *GENDER*: A dummy variable equals to 1 if the respondent is a female.
- *AGE*: Age of the respondent in years.
- *HHsize*: Size of the respondent's household.
- *Sh_child*: Share of children (<15) in the household.
- *INC*: Income quintile indicator, ranging from 1 (lowest income quintile) to 5 (highest income quintile).

We use all the common variables across both surveys as regressors, adding non-linearities in the effects of age and household size. We are not very much concerned about endogeneity in our setting because we are interested in obtaining the overall impact of the variables to maximize the efficiency of the imputation.

The above regression is estimated within the Multiple Imputation MI methods framework, by using ordered logit since the dependent variable Raw Score of the FIES is a discrete ordinal one. Table 2 provides the estimation results for the original survey data in the FAO dataset. Most independent variables' coefficients were statistically significant and have the expected signs.

Table 2: Ordered logit estimation results for the determinants of food security score in survey 1

Variables	Raw_score
Gender (Female=1)	-0.0245
	(0.118)
Age	0.0537***
	(0.0167)
Age squared	-0.000592***
	(0.000178)
Basic education (=1)	0.913***
	(0.133)
Household size	-0.272**
	(0.122)
Household size squared	0.0171
	(0.0111)
Share of children	0.842***

	(0.316)
Quintile Income	-0.699***
	(0.0504)
/cut1	-2.455***
	(0.533)
/cut2	-1.909***
	(0.531)
/cut3	-1.324**
	(0.528)
/cut4	-0.798
	(0.527)
/cut5	-0.328
	(0.526)
/cut6	0.195
	(0.526)
/cut7	0.839
	(0.528)
/cut8	1.782***
	(0.536)
LR chi2(8) = 381.72	Prob > chi2 = 0.0000
Log likelihood = -1800.1451	Pseudo R2 = 0.0959
Observations 984	

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Estimation implemented in Stata 16.

The common variables used above are selected from survey 2 and standardized using the variables' means and standard errors from survey 1 (as detailed above). Then the Raw Score missing data in survey 2 is imputed by using MI methods with the determinants of the food security raw score model detailed previously.

The prevalence rates of moderate and severe Food Insecurity (FI) are then calculated in the 2024 Arab Barometer survey based on the imputed Raw Scores, using the randomized cluster survey design. Moderate to severe food insecurity is obtained for individuals with Raw Score equal or above 4, while severe food insecurity is for individuals with a Raw Score equal or above 7 (consistent with the recommended thresholds in Sheikomar et al., 2021).

Table 3 provides the imputation-based national estimates of food insecurity in Lebanon for 2024, compared to the actual rates computed in 2023 based on the FAO FIES survey. Notice that our estimates are quite close to the actual rates, with Moderate to Severe FI rates ranging closer to those of 2023.

Table 3: Predicted food insecurity rates

	Survey 1 - FAO (2023)	Survey 2 – Arab Barometer (2024)
Prevalence rates of food insecurity among the adult population %	Actual	Imputation-based estimates
Severe	15.08 (1.23)	13.59 (1.42)
Moderate to Severe	39.60 (1.72)	36.78 (2.02)

Note: Standard errors are in parentheses. We use 50 simulations for MI methods imputation. Imputations implemented in Stata 16.

3.2. Imputation of food insecurity incidence at the sub-national level

As discussed above, we are also very interested in computing subnational food insecurity prevalence rate estimates across governorates, as this allows for better design of policies targeting food insecurity and also allows to track the evolution of food insecurity at the subnational level. Since survey 1 (FAO 2023) did not collect information on governorates while survey 2 (Arab Barometer 2024) did, we can compute the prevalence rates of Severe and Moderate to Severe FI in 2024 in survey 2 based on the imputed rates and the cluster design. Table 4 details these rates across the country's eight governorates. These provide prevalence estimates until April 2024 (date when survey 2 was completed).

Table 4: Prevalence rates of food insecurity among adults 18+ at the sub-national level, %

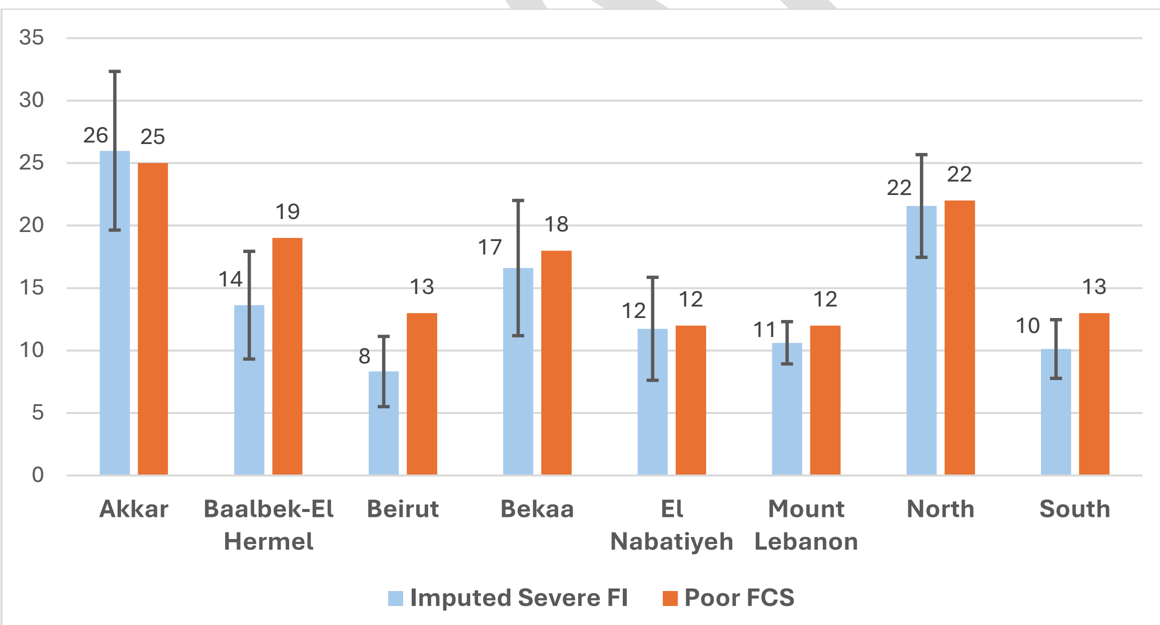
	Severe		**Moderate to severe**	
Governorate	Mean	Std. err.	Mean	Std. err.
Akkar	25.99	6.35	59.29	6.61
Baalbek-El Hermel	13.63	4.31	37.66	6.56
Beirut	8.32	2.81	26.58	4.28
Bekaa	16.60	5.41	41.22	8.21
El Nabatiyeh	11.74	4.12	33.64	6.19
Mount Lebanon	10.62	1.69	30.60	2.72
North	21.57	4.11	52.89	4.82
South	10.12	2.35	31.30	3.75

Source: Based on authors' computations.

In order to evaluate the validity of these estimates, we rely on recent food security data collected in the country which could provide an indirect benchmark for the imputed estimates. The World Food Programme (WFP) and the FAO regularly collect in Lebanon household food consumption data, and compute a Food Consumption Score (FCS) indicator (FAO and WFP, 2024). This indicator is a composite score based on households' dietary diversity, food consumption frequency, and relative nutritional value of different food groups. The FCS is calculated by asking how often households consume food items from the 8 different food groups (plus condiments) during a 7-day reference period.

Although the FCS is calculated at the household level, it provides an indirect benchmark for evaluating the FI estimates, as food consumption is very closely linked to experienced food security. Figure 1 details the mean percentage of poor and borderline food consumption scores (FCS) by governorate in the period November 2023 – August 2024, which are a good proxy for severe food insecurity. These percentages are plotted against the imputed severe FI prevalence rates by governorate. We find that for 6 out of the 8 governorates the imputed rates and the poor and borderline FCS rates are very close, with the difference in the other two governorates falling within the estimated range of imputed FI. It is to be noted that both indicators are computed in a different way (the FI rate is opinion-based and is for adult individuals 18+ while the FCS is for household consumption data at the household level). Yet the fact that both indicators have close values within the same time period provides a good external validity check for our subnational estimates.

Figure 1: External validity check for subnational FI estimates
Comparison of the imputed severe food insecurity prevalence rates (%) with the percentage of poor and borderline food consumption scores (FCS) by governorate, November 2023–August 2024



Source for FCS data: FAO and WFP (2024). Standard error plots of imputed FI estimates in line intervals.

4. Discussion

This paper has shown that imputed food insecurity prevalence indicators can be obtained in a robust manner with few restrictions and with a few explanatory variables. In addition to this, the estimation of subnational food insecurity prevalence indicators, especially in conflict-stricken countries such as Lebanon, is of central importance for policymakers implementing programs targeting food insecurity. The availability of sub-national food insecurity measures at the regional level provides a crucial opportunity for better design and targeting of on-going food aid programs, as was shown in recent research that used regional variations in recipient characteristics to refine food insecurity targeting programs (Chaaban et al., 2018, Hatzenbuehler and Mavrotas, 2021).

Our cross-survey imputation methodology for food insecurity incidence can also be used to regularly update food security estimates in resource-constrained countries that cannot implement the FIES module on regular basis. Our results show that it is possible with a few common variables and nationally representative surveys to impute food security estimates in a relatively straightforward manner.

With this in mind, it is important for future research to conduct more cross-validity checks for imputed estimates, by realizing within-survey imputations in large datasets (often available for large countries), in order to test the stability and representativeness of imputed food insecurity estimates. Future research could also incorporate other food insecurity indicators such as undernourishment and food consumption in the set of indicators to be imputed across surveys. This would allow a richer analysis and refinement of national, regional and global estimates of indicators related to SDG Target 2.1.

References

Ahmed Jubayer, Saiful Islam, Abira Nowar, Md. Moniruzzaman Nayan, Md. Hafizul Islam. (2023) Validity of Food insecurity experience scale (FIES) for use in rural Bangladesh and prevalence and determinants of household food insecurity: An analysis of data from Bangladesh integrated household survey (BIHS) 2018-2019, *Heliyon*, Volume 9, Issue 6, e17378.

Arab Barometer Wave VIII Questionnaire and Technical Report, 2023-2024.

Ballard, T.J., Kepple, A.W. & Cafiero, C. (2013). The food insecurity experience scale: developing a global standard for monitoring hunger worldwide. Technical Paper. Rome, FAO.

Bartfeld, Judith, and Madeline Reed-Jones. (2024). " Can the American Community Survey Provide New Insight into Household Food Security? An Illustration of Cross-Survey Multiple Imputation." *Applied Economic Perspectives and Policy* 46(4): 1627–1645.

Beltramo, T., Dang, H. A., Sarr, I., & Verme, P. (2024). Estimating poverty among refugee populations: a cross-survey imputation exercise for Chad. *Oxford Development Studies*, 52(1), 94–113.

Chaaban, J., Ghattas, H., Irani, A., & Thomas, A. (2018). Targeting mechanisms for cash transfers using regional aggregates. *Food Security, 10*, 457–472.

Dang, Hai-Anh H., Peter F. Lanjouw, Umar Serajuddin, (2017) Updating poverty estimates in the absence of regular and comparable consumption data: methods and illustration with reference to a middle-income country, *Oxford Economic Papers*, Volume 69, Issue 4, October, Pages 939–962.

Dang, HA.H., Verme, P. (2023) Estimating poverty for refugees in data-scarce contexts: an application of cross-survey imputation*. J Popul Econ* 36, 653–679.

FAO, IFAD, UNICEF, WFP and WHO. 2024. *The State of Food Security and Nutrition in the World 2024 – Financing to end hunger, food insecurity and malnutrition in all its forms.* Rome. https://doi.org/10.4060/cd1254en

FAO Statistics Division, Lebanon - Food Insecurity Experience Scale (FIES). July 18, 2024. https://microdata.fao.org/index.php/catalog/2610/data-dictionary/F5

FAO and WFP. 2024. Special Report – 2024 FAO/WFP Crop and Food Security Assessment Mission (CFSAM) to the Lebanese Republic. November 2024. CFSAMs Special Reports, 03/2024. Rome. https://doi.org/10.4060/cd3356en

Hatzenbuehler, P.L., Mavrotas, G. Combining household and price data to target food marketing interventions in Nigeria. *Food Sec.* 13, 493–505 (2021).

Mahmood, T., Kumar, R., Ali, T. M., Naeem, N., & Pongpanich, S. (2023). Determinants of the food insecurity at household level in Pakistan: A multilevel model approach. *PloS one*, 18(10), e0291343.

Raghunathan, T. E., J. M. Lepkowski, J. Van Hoewyk, and P. Solenberger. 2001. A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodology 27: 85–95

Saint Ville, A., Po, J.Y., Sen, A. *et al.* (2019) Food security and the Food Insecurity Experience Scale (FIES): ensuring progress by 2030. *Food Sec.* 11, 483–491.

Sheikomar OB, Dean W, Ghattas H, et al. (2021) Validity of the Food Insecurity Experience Scale (FIES) for Use in League of Arab States (LAS) and Characteristics of Food Insecure Individuals by the Human Development Index (HDI). Curr Dev Nutr 5.

van Buuren, S. 2007. Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research* 16: 219–242. https://doi.org/10.1177/0962280206074463.

Wooldridge, J.M. Econometric Analysis of Cross Section and Panel Data. The MIT Press. Cambridge, Massachusetts. 2010.