WORKING PAPERS

N° 18-896

Februray 2018

“Targeting mechanisms for cash transfers using regional aggregates”

Jad Chaaban, Hala Ghattas, Alexandra Irani and Alban Thomas

Toulouse School of Economics

Food Security
https://doi.org/10.1007/s12571-018-0768-5

ORIGINAL PAPER

Targeting mechanisms for cash transfers using regional aggregates

Jad Chaaban[1] · **Hala Ghattas**[2] · **Alexandra Irani**[1] · **Alban Thomas**[3]

Received: 8 August 2016 / Accepted: 31 January 2018

Abstract
We propose an empirical method for improving food assistance scoring and targeting, which minimizes under-coverage and leakage of food and cash assistance programs. The empirical strategy relies on a joint econometric estimation of food insecurity and economic vulnerability indicators at the household level, using data-driven instead of predetermined quantiles. We apply the method to recent micro data on Syrian refugees in Lebanon, to explore how regional and community-based aggregates can improve the targeting effectiveness of aid programs, notably food aid by the World Food Program in Lebanon. Our results confirm that using regional aggregates are useful for augmenting the Balanced Poverty Accuracy Criterion, and our method performs much better than the current policy in terms of targeting effectiveness and accuracy for economically vulnerable households.

Keywords Targeting · Food security · Economic vulnerability · Food aid · Refugees

1 Introduction

A major challenge of food policy consists of targeting, in a cost-effective way, poor households that may be both food insecure and economically vulnerable. An efficient targeting would, in particular, succeed in limiting under-coverage and leakage of food cash assistance programs. When dealing with households' status of food security and economic welfare as the target of aid policies, it is important to distinguish between the concept of vulnerability and the one of insecurity. In general, (economic) vulnerability is a measure of the risk a household may drop below some welfare measure (usually, the poverty line), while (food) insecurity indicates a current status of the household, regarding food access and consumption (see Dercon 2006).

The main objective of this paper is to derive an empirical method for improving food assistance scoring and targeting systems in situations of budget limitations, which can be used by decision makers and analysts. More precisely, we suggest a data-driven method for targeting food insecure households, using both household and community-level indicators of food security and welfare. Our empirical strategy explores how different levels of information on households and administrative (district) average characteristics can be used to reduce under-coverage and leakage of food cash assistance programs, in order to increase the performance of the food aid system in targeting poor households. Such empirical strategy relies on a robust, joint estimation of food insecurity and economic vulnerability indicators at the household level, using data-driven instead of predetermined quantiles.

We consider the World Food Programme's (WFP) food assistance system in Lebanon and analyze its scoring and targeting system for Syrian refugees in that country. Six years into the Syrian crisis, Lebanon hosts just over 1 million Syrian refugees, who are registered with the office of the United Nations High Commissioner for Refugees (UNHCR), about 50% of whom are under 15 years of age (WFP, 2016). Sequential surveys conducted by United Nations agencies have consistently found a large share of Syrian refugees in

✉ Alban Thomas
Alban.Thomas@inra.fr

Jad Chaaban
jc11@aub.edu.lb

Hala Ghattas
hg15@aub.edu.lb

Q1/ Q2
[1] Department of Agriculture, American University of Beirut, Beirut, Lebanon

[2] Center for Research on Population and Health, American University of Beirut, Beirut, Lebanon

[3] Toulouse School of Economics, I.N.R.A, University of Toulouse, 31015 Toulouse Cedex 6, France

Lebanon to be living below the poverty line (71% in 2016), to experience some level of food insecurity and to be adopting coping strategies that involve the depletion of assets and incurring of debt to cover food, health and rental expenses (WFP, 2016). In response, various humanitarian agencies have established multi-purpose cash assistance programs, in addition to cash-based food assistance and in-kind food assistance targeting vulnerable households.

We use original and unique data for our statistical and econometric analysis on micro data collected from refugee households, to evaluate the determinants of food insecurity and economic vulnerability. We investigate the empirical relationship between food insecurity and economic vulnerability at the household level, by estimating a structural system of simultaneous ordered Probit equations for both indicators.

The empirical methodology is based on a multi indicator system (proxy means, a point or a scoring system), which contains observable (and easy to verify) household characteristics. This system, devised through statistical analysis, allows prediction of food insecurity scores for the already sampled and remaining refugee population within a well-defined margin of error that reduces targeting inclusion and exclusion errors. This would prove useful in first analyzing the current food security status and the targeting mechanism of refugees using all available data. Second, introducing a model to identify refugee households' food security status allowing better targeting of the most vulnerable households, with observable exogenous indicators most closely correlated with food insecurity. Last, identifying key observable indicators, which could help in-depth monitoring through follow-up visits.

As described in a survey of food security by Barrett (2002), "effective targeting is fundamental to food aid policy (FAP) design and evaluation, particularly in today's era of shrinking FAP budgets as a proportion of government spending or gross domestic product (p. 58)". The purpose of cost-effective targeting is to reduce leakage to unintended beneficiaries and to maximize the proportion of poor households effectively participating in the program (Borjas 2004). The literature on targeting categories of populations is mostly dedicated to poverty alleviation policies and access to natural resources and energy (water utilities, etc.). Many papers have examined the performance of food stamp programs (in developing countries and mostly in the United States) when the poverty status of households is costly to verify and administrative costs may reduce the amount of resources allocated to the poverty intervention (Wilde and Nord 2005; Barrett 2002; Besley and Kanbur 1993). When reliable data on household income are difficult to obtain, proxy means test can be considered from a variety of instruments observed with low cost, assumed correlated with household welfare and difficult to modify by households (Sen 1995; Glewwe 1992). Moreover, even though poor household registration and regular re-certification procedures may increase administrative costs dramatically in the short run, their impact on improving targeting is likely to be visible in the longer run.

Benfield (2007) considered indicators that may improve the performance of a food-stamp policy in Jamaica through better targeting by minimizing Type-I and Type-II errors. The paper confirms that a policy based on additional indicators, e.g., on housing conditions and durable goods, may perform better in targeting the poor. Another stream in the literature concerns spatial patterns of poverty and implications for food policies of the spatial distribution of poor households in poor/non-poor geographical areas (Minot and Baulch 2005; Kam et al. 2005; Amarasinghe et al. 2005; Elbers et al. 2007; Jaynes et al. 2001; Agostini and Brown 2011). Indeed, among the many factors that explain that targeting may not be cost-effective, there is the possibility that the target group is a spatially dispersed proportion of the general population (Barrett 2002). Barrett (2002) distinguishes between ICAT (Indicator-Contingent Administrative Targeting) and self-targeting programs, the former relying on screening based on various food security indicators (including income and nutritional status) to determine program eligibility of households. By contrast, self-targeting is designed so that only intended beneficiaries participate on a voluntary basis, with asymmetric information issues partly solved through some cost component or a reduction in quality of the good or service. Some empirical research suggests that easy-to-collect indicators such as dependency ratio, rooms per capita, etc. can be fairly effective in identifying food-insecure households with some degree of cost savings (Haddad et al. 1997; Chung et al. 1997; Lipton and Ravallion 1995). See, e.g., Jaynes et al. (2001) for shortcomings and caveats of such indicator-based targeting strategies.

The rest of the paper is organized as follows: Section 2 details the materials and methods (theoretical and empirical methodologies) we employed throughout the analysis. This includes the description of the datasets and variables used in our empirical application, and the empirical strategy. In particular, the specification of the simultaneous ordered Probit system of equations for food insecurity and economic vulnerability is presented in detail. Section 3 presents the estimation results from the structural system of equations, and the performance of our empirical method in terms of targeting accuracy and effectiveness. Such performance analysis entails a comparison of coverage and under-coverage of poor households, leakage and targeting differential associated with various specifications, including community-level and regional aggregates. Finally, concluding remarks are in Section 4.

2 Materials and methods

2.1 Theoretical model

As discussed above, better targeting of poor households may help to attain the objective of poverty alleviation at lower costs, but policy makers are faced with the prohibitive costs associated with the identification of households below the poverty line. The trade-off therefore involves monitoring costs on the one hand, and policy leakage (to non-poor households) on the other, with the general objective of increasing welfare of food insecure and/or economically vulnerable households. To formalize the terms of the trade-off, we introduce a simple model that represents the targeting issue in a stylized food policy.

Consider a food-aid planner whose objective is to target poor households with a given cash transfer denoted A (per HH), from a total population of size (N, known). The unit administrative cost of transferring A to a household is denoted τ, the proportion of poor (non-poor) households in the total population is β_P (respectively, β_{NP}). Let $p_1(\alpha)$ and $p_2(\alpha)$ respectively denote the probability of under-coverage (not targeting a poor household) and leakage (targeting a non-poor household), which depend on the effort of collecting and treating information on households, denoted α. The cost of effort is $F(\alpha)$ and we assume

$$\frac{dp_1(\alpha)}{d\alpha} \leq 0, \frac{dp_2(\alpha)}{d\alpha} \leq 0, \frac{dF(\alpha)}{d\alpha} > 0. \tag{1}$$

For a given level of effort α, the social planner wishes to obtain a particular level of targeting and, therefore, a given level of social welfare as an outcome. We assume the social planner solves this problem by determining both the optimal for effort (α) and the level of cash transfer (A).

There are several options the social planner can, in theory, choose from to determine the optimal food policy. A first option for the social planner is to define a proportion of population to be assisted, say, $\underline{\beta}_P$, and then solves for the optimal level of targeting effort jointly with the unit level of cash transfer. Note that in such a case, targeting is meaningful, at least in the sense defined above, that is, identifying poor households to receive aid. When such an option is selected, households are simply ranked by increasing order of income or wealth and the first $N\underline{\beta}_P$ receive assistance.

A second possibility is that the social planner determines the optimum level of aid per household (i.e., the monetary level of cash transfer) by dividing total budget by the number of poor households,

$$A^* = \frac{(B - F(\alpha))}{N\tau\beta_P}, \tag{2}$$

i.e., accounting for administrative costs. In this case, the minimization problem becomes, after substituting for A^* and using $\beta_{NP} = 1 - \beta_P$,

$$max_\alpha\ C = [B - F(\alpha)] \times \left\{ [1 - p_1(\alpha)] - \frac{1-\beta_P}{\beta_P} p_2(\alpha) \right\}, \tag{3}$$

which gives

$$\begin{aligned} &\frac{\partial F(\alpha)}{\partial \alpha} \times \left\{ [1 - p_1(\alpha)] - \frac{1-\beta_P}{\beta_P} p_2(\alpha) \right\} \\ &= -[B - F(\alpha)] \times \left\{ \frac{dp_1(\alpha)}{d\alpha} + \frac{1-\beta_P}{\beta_P} \times \frac{dp_2(\alpha)}{d\alpha} \right\} \geq 0, \end{aligned} \tag{4}$$

which implies that $\{[1 - p_1(\alpha)] - \frac{1-\beta_P}{\beta_P} p_2(\alpha)\} \geq 0$, provided the cost of effort does not exceed the initial budget, i.e., $[B - F(\alpha)] \geq 0$. This second option is obviously not relevant when the objective is to help poor households reach some poverty line level. With such a policy, some assisted households could become better off than non-assisted ones, which would introduce distortions in the distribution of households.

A third possibility is that the social planner first determines exogenously the cash transfer that corresponds to a minimum level of food expenditure to ensure food security. For example, some minimum expenditure level can be computed given local prices, based on requirements for food and nutrient intake provided by international standards. More precisely, the social planner would equate direct and indirect utility levels

$$V\left(p, \underline{y} + A\right) = U(FS), \tag{5}$$

where $V(.)$ and $U(.)$ are indirect (Hicksian) and direct (Marshallian) utility functions respectively, p is the vector of market prices faced by households, $\underline{y}$ is exogenous income (not depending on cash transfers) and FS is the level of food expenditure to guarantee food security. Solving for cash transfer in the equation above allows the social planner to determine the optimal level of transfer to reach food security, denoted A^*, given available income and price levels. Then, in a second stage, the social planner solves the following program:

$$\begin{aligned} max_\alpha\ C = A^* \times \tau \times N \\ \times \{\beta_P[1 - p_1(\alpha)] - \beta_{NP} p_2(\alpha)\} - F(\alpha), \end{aligned} \tag{6}$$

such that $C \leq B$ (total budget available). We have

$$\frac{\partial F(\alpha)}{\partial \alpha} = -\left(A^{*}\tau N\beta_P\right) \times \left\{\frac{dp_1(\alpha)}{d\alpha} + \frac{1-\beta_P}{\beta_P} \times \frac{dp_2(\alpha)}{d\alpha}\right\}. \quad (7)$$

In other words, the optimal level of effort to obtain information about food insecurity through, e.g., additional surveys, is determined when the marginal cost of such effort (on the LHS) equals the marginal benefits of effort in terms of better targeting (on the RHS). Such benefits are a weighted sum of marginal effects (with respect to targeting effort) in under-coverage and leakage probabilities, where the weight is the ratio of the non-poor over the poor proportion of households in the total population. Under such constraint, given a predetermined level of cash transfer per household, the policy maker determines the optimal level of targeting that ultimately determines the proportion of poor households to be provided with assistance.

This last policy corresponds to the one employed in practice in our application, and we will consider it in the rest of the paper.

When the level of effort tends to infinity, we expect both policy under-coverage and leakage to tend to zero, so that the total cost of the targeting policy would converge to

$$\begin{aligned} C &= \left(A^{*}\tau N\beta_P\right) \times \left\{1 + \frac{1-\beta_P}{\beta_P} \times 0\right\} - F(\alpha) \\ &= \left(A^{*}\tau N\beta_P\right) - F(\alpha). \end{aligned} \quad (8)$$

This would violate the condition that $C \leq B$ if $F(\alpha)$ is large enough. However, a trade-off can be determined by solving the above problem, for a limited budget and expected gains from targeting. In practice, it is essential to be able to identify effective gains from targeting poor households, and compare them with the cost of effort associated with data collection on households. This is achieved by sampling over the target population to estimate the proportion of poor households therein, providing an estimate for β_P. When the sample includes households benefiting from cash transfers as well, then the probabilities of under-coverage and leakage can also be estimated. The empirical analysis of the present paper proposes a system of targeting equations that serve such a purpose by illustrating the way increasing information on households can increase the performance of targeting policies, by reducing probabilities $p_1(\alpha)$ and $p_2(\alpha)$. To make the connection between the above model and our application, we assume a direct relationship between targeting effort and information collected on the population of households. As information increases through (costly) repeated surveys, the precision on coverage and leakage probabilities also increase in a way limited by the trade-off discussed, until the marginal benefit of targeting effort equals its marginal cost.

2.2 Empirical methodology

In this paper we are interested in analyzing the targeting effectiveness of cash transfers with respect to two variables: food insecurity and economic vulnerability. As will be shown below, both dependent variables are discrete ordered, and they are interlinked. This requires a specific empirical estimation strategy, which is detailed here.

Instead of specifying a reduced-form system of equations in which both indicators would appear dependent upon the same set of socio-economic explanatory variables, we consider instead a structural specification where food insecurity depends on economic vulnerability as well. The argument behind such specification is that it provides a simple way to disentangle the effect of socio-economic drivers of food insecurity from the ones (possible in common) that condition welfare and economic vulnerability. Moreover, with such representation, it is straightforward to simulate the impact of a change in income following, e.g., a decrease in the cash transfer, and its impact on the food insecurity indicator. In any case, the recursive representation at the household level is consistent with the evidence that, even though identical socio-economic determinants may jointly explain food insecurity and economic vulnerability, the former is also determined by the level of economic welfare.

Let i, $i = 1,2,\ldots, N$, denote the household index and consider the general simultaneous-equation model:

$$\left\{y_{1i}^{*} = \delta_1 + x_{1i}\beta_1 + u_{1i}, y_{2i}^{*} = \delta_2 + \gamma y_{1i}^{*} + x_{2i}\beta_2 + u_{2i}, \right. \quad (9)$$

where y_{1i}^{*} and y_{2i}^{*} are two (continuous) latent variables that can be defined as measures of welfare related to income and food respectively. They are associated with two observed simultaneous welfare levels (respectively, food insecurity and economic vulnerability) and are assumed to be positive when corresponding levels are observed. Vectors of explanatory variables x_{1i} and x_{2i} may have some common components, u_{1i} and u_{2i} are random variables with a correlation coefficient between denoted ρ (assumed constant). We assume the following exogeneity restrictions apply: $E(x_{1i}u_{1i}) = E(x_{2i}u_{2i}) = 0, \forall i.$.

Both latent variables typically lie in the real line, as both food insecurity and economic vulnerability levels may be normalized to correspond to a set of non-overlapping intervals with negative and positive values. Let $\left\{S_j^k = \left[c_{j-1}^k, c_j^k\right], j = 1, 2, \ldots, J_k; k = 1, 2\right\}$ denote such non-overlapping sets, with $\cup_j S_j^k = R, \forall k = 1, 2, c_0^k = -\infty$, $c_{J_K}^k = \infty, \forall k$, and $c_{j-1}^k \leq c_j^k, \forall k, \forall j.$

JrnlID 12571_ArtID 768_Proof# 1 - 02/02/2018

In the dataset we observe ordered dependent variables: $y_{1i} = 1$ if $y_{1i}^{*} \in S_j^1$ and $y_{2i} = 1$ if $y_{2i}^{*} \in S_j^2$, $j = 1, 2, \ldots, J_1, k = 1, 2, \ldots, J_2$.

From the structural system of Eq. (1), we have:

$$
\begin{aligned}
&Prob\left(y_{1i}^{*} \in S_j^1, y_{2i}^{*} \in S_k^1\right) = Prob(y_{1i} = j, y_{2i} = k)\\
&\quad = Prob\left(c_{j-1}^1 \le y_{1i}^{*} < c_j^1, c_{k-1}^2 \le y_{2i}^{*} < c_k^2\right)\\
&= \Phi_2\left[c_j^1 - \delta_1 - x_{1i}\beta_1, \theta\left(c_k^2 - \gamma\delta_1 - \gamma x_{1i}\beta_1 - \delta_2 - x_{2i}\beta_2\right), \underline{\rho}\right]\\
&-\Phi_2\left[c_{j-1}^1 - \delta_1 - x_{1i}\beta_1, \theta\left(c_k^2 - \gamma\delta_1 - \gamma x_{1i}\beta_1 - \delta_2 - x_{2i}\beta_2\right), \underline{\rho}\right]\\
&-\Phi_2\left[c_j^1 - \delta_1 - x_{1i}\beta_1, \theta\left(c_{k-1}^2 - \gamma\delta_1 - \gamma x_{1i}\beta_1 - \delta_2 - x_{2i}\beta_2\right), \underline{\rho}\right]\\
&+\Phi_2\left[c_{j-1}^1 - \delta_1 - x_{1i}\beta_1, \theta\left(c_{k-1}^2 - \gamma\delta_1 - \gamma x_{1i}\beta_1 - \delta_2 - x_{2i}\beta_2\right), \underline{\rho}\right],
\end{aligned}
\tag{10}
$$

where $\Phi_2(\cdot, \cdot, \cdot)$ is the bivariate standard normal cumulative distribution function, and $\theta = (1 + 2\gamma\rho + \gamma^2)^{-1/2}$, $\underline{\rho} = \theta(\gamma + \rho)$ ..

The expression (2) can be evaluated for any pair of outcomes (j, k) and all contributions of the sort are used to construct the log-likelihood of the sample, to obtain consistent Maximum Likelihood estimates of the bivariate ordered Probit (see Sajaia 2007). $J_1 + J_2 - 1$ cut off values $\left(c_j^k\right)$ are estimated together with parameters ($\beta_1, \beta_2, \gamma, \rho$), but intercept terms δ_1 and δ_2 are not identified (in fact, cut offs are identified up to a constant term). Parameters in (1) are only identified if we impose exclusion restrictions, that is, at least one variable in x_{1i} should be excluded from x_{2i}. An interesting candidate for such exclusion is an exogenous variable that determines economic vulnerability but not food insecurity (such as the particular assets possessed by the household, as in the application).

Our model specification has economic vulnerability y_1^{*} as an explanatory variable in the equation for food insecurity and this variable is endogenous by construction. If error terms u_1 and u_2 are correlated ($\rho \neq 0$), it implies that y_{1i}^{*} is correlated with u_{2i}, and the second equation in the system (1) cannot be estimated independently. In our empirical analysis of joint estimation of food insecurity and economic vulnerability, this endogeneity issue is essential to avoid simultaneity bias in parameter estimates.

To test for the endogeneity of y_{1i}^{*} in the equation for y_{2i}^{*}, we estimate the structural system of equations by bivariate (ordered) Probit and the Full Information Maximum Likelihood (FIML) method. We then use a Wald Test of $\gamma = 0$ in the second one, see Sajaia (2007). Note that we do not consider, for the sake of space limitation, an alternative estimation method that would consider a bivariate ordered Probit applied to the reduced form of the system of equations. Even though such specification could be considered to provide us with consistent parameter estimates (as long as exogeneity of y_2^{*} in the sense defined above is rejected), we are able to obtain structural parameter estimates directly by FIML with the bivariate ordered Probit procedure.

The log-likelihood function over N observations is

$$
log\ L = \sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{k=1}^{K} I(y_{i1} = j, y_{i2} = k) log\ Prob(y_{i1} = j, y_{i2} = k) \tag{11}
$$

if observations are identically and independently distributed. This may not be the case, in particular when unobserved random effects come in addition to random errors u_1 and u_2. A possibility in this case is to specify the joint distribution of such effects and to integrate them out from the following log-likelihood:

$$
log\ L(\pi) = \int f(y|x, u, \pi)\phi(u|\mu_u, \Sigma_u)du, \tag{12}
$$

where π is the vector of structural parameters, u is the vector of random effects with joint distribution defined by the density function $\phi(u|\ \mu_u, \Sigma_u)$. Such integral can be approximated by Gauss-Hermite quadrature, see, e.g., Rabe-Hesketh et al. (2005) and Skrondal and Rabe-Hesketh (2004). We now describe the main data sets used by WFP for targeting, as well as the ones we use in the paper to construct our food security and welfare indicators.

2.3 Main datasets used

To deal with an ever-increasing number of refugees reaching Lebanon from the Syrian border, international organization such as the WFP, the United Nations Children's Fund (UNICEF), and the UNHCR in Lebanon initiated the yearly Vulnerability Assessment on Syrian Refugees (VASyR), for programmatic purposes (UNHCR, UNICEF and WFP 2015). It aims to construct eligibility criteria for targeting beneficiaries using individual data from refugee surveys containing both observed and self-reported variables. Based on VASyR, WFP in Lebanon developed a vulnerability scoring system dedicated to household targeting (see Drummond et al. 2015).

Household vulnerability is generally defined as the likelihood of a household to not cover basic needs of all members without engaging in irreversible coping strategies due to a lack of financial resources (World Food Program 2016). In this regard, it is a measure of the risk of moving to a less favorable status, and it needs to be distinguished from the food security status of the household (a current state that may be affected by factors external or internal to the household). This distinction is important to recall here, because the “vulnerability“assessment of VASyR includes both economic vulnerability and food security dimensions, as discussed below, whereas our empirical analysis is directed towards two indicators only: food insecurity and economic vulnerability. Eligibility criteria in the formula used by the WFP include indicators of food security and economic vulnerability, as well as self-reported coverage of household basic needs. Eight sector-specific vulnerabilities are included in the WFP vulnerability scoring system, including food security, economic vulnerability, education, health, non-food items, protection, shelter, etc. A household is classified into one of four vulnerability categories according to each of these eight sectors, and the sector scores are then summed to produce a global vulnerability score comprised of five vulnerability categories: low, mild, moderate, high and severe.

The main datasets used in our paper are detailed below:

2.3.1 VASyR 2015

The VASyR 2015 (Vulnerability Assessment on Syrian Refugees) is a nationally representative two-stage cluster survey of Syrian refugees living in Lebanon (see World Food Program 2016), conducted in May–June 2015 and includes around 4100 households for an estimated 21,300 individuals. VASyR 2015 was used as the main data set to estimate and test improved food security and economic vulnerability indices. The data set was subject to data cleaning (duplicate observations, multiple heads of households, etc.) for a final sample size of 3850 households.

2.3.2 ProGres

The Profile Global Registration System (ProGres) is the main global database used by the UNHCR and the data provided include all registered refugees in Lebanon (about 1.05 million individuals in October 2015). Although valuable because of its size and the inclusion of key socioeconomic characteristics on refugees, the database does not contain variables measuring welfare. Although registration is only voluntary, due to immigration rules in hosting countries and UNHCR procedures and incentives to register, most refugees would register at some point in time.

The household is considered the main unit of measurement for the VASyR survey. It is defined by WFP as a group of people, who routinely eat out of the same pot, live in the same compound (or physical location), and share the same budget, managed by the head of household. In contrast, the “case” is used by UNHCR to register refugees in the ProGres data base and is defined as: “A processing unit similar to a family headed by a Principal Applicant. It comprises (biological and non-biological) sons and daughters up to the age 18 (or 21) years, but also includes first degree family members emotionally and/or economically dependent and for whom a living on their own and whose ability to function independently in society/in the community and/or to pursue an occupation is not granted, and/or who require assistance from a caregiver.”[1]

WFP’s targeting process starts from blanket coverage and then focuses on identifying and removing households and individuals who do not need food assistance according to a certain vulnerability threshold. UNHCR’s targeting of its unconditional cash transfers moves in the opposite direction by focusing on identifying the refugees who are most in need of economic assistance, and then expanding coverage as resources allow and as needy cases are identified.

The WFP vulnerability scoring system is then applied on the households visited using the common multi-agency questionnaire (over 90,000 households) and used for targeting purposes: households are excluded from assistance if they fall within the better off vulnerability categories (low, mild, moderate), and according to a combination of economic considerations and Multi-functional Team (MFT) revisions.

There are obvious limitations of the WFP vulnerability scoring system, including lengthy and expensive household visits and the fact that it includes over 50 variables, some of which are duplicated within the score, rendering it difficult to use as a desk formula. Finally, the formula includes both input and output variables, which would lead to endogeneity problems in a desk formula. For example, some of the variables used to calculate the WFP vulnerability score include the food consumption score and a coping strategies index both of which can be considered outcome variables for food security. The score also includes input variables such as dependency ratio, education, gender of the head of household, members with a disability amongst other input variables.

Appendix 1 details the derivation of the food security indicator, obtained through a fully data-driven procedure directly from individual surveys. The major advantage of our procedure is that empirical quantiles allow determining empirical values that do not depend on external standards that may be inconsistent with local conditions. Moreover, in designing our indicators, we carefully exclude variables likely to reflect decisions from households that may depend on other explanatory variables (the endogeneity problem, see below).

Table 1 displays the proportion of households classified into five groups of vulnerability (regarding food insecurity), according to the index derived in Appendix 1.

Defining food insecurity by a gradient above 4, the proportion of food insecure households in the VASyR2015 sample is around 61%.

As for economic vulnerability, we consider a welfare-dependent variable that proxies economic vulnerability, on a set of independent variables that are thought to determine the

[1] http://cega.berkeley.edu/assets/miscellaneous_files/35-ABCA_-Targeting_and_Welfare.pdf

Table 1 Proportion of households in VASyR 2015 classified into categories of vulnerability to food insecurity

Lowest to highest vulnerability to food insecurity	Proportion of households
1 – lowest	13.3%
2	10.6%
3	14.7%
4	33.2%
5 – highest	28.2%

Table 2 Economic vulnerability quantiles

Economic vulnerability quantile	Value (USD/month/head)	Observation s in sample
10%	41.66	269
20%	55.33	328
30%	67.50	458
40%	78.33	500
50%	90.46	633
60%	104.62	552
70%	121.68	511
80%	146.80	337
90%	195.38	183
Average	112.47	3771

variation in the welfare aggregate. We adopt per capita expenditure as our welfare aggregate. Economic vulnerability is therefore measured through monthly expenditure per capita in USD. The expenditure aggregate is constructed by summing up 18 self-reported expenditure items from the VASyR questionnaire with a recall period of 30 days. The household aggregate is then divided over the number of household members. Following the approach used by UNHCR, the upper limit of the per capita expenditure aggregate was restricted to 250 USD to exclude most outliers. As for the food insecurity indicator, we construct an ordinal variable based on bootstrapped quantiles of the monthly expenditure per capita, with values from 1 to 9. Quantiles of economic vulnerability are reported in Table 2, as monthly values in USD per capita, together with the corresponding number of households in the sample.

2.4 Model specification

Two targeting models are detailed below, ranging from the inclusion of ProGres only variables to appending community level indicators. Recall that in all model specifications, the equation for food insecurity contains the endogenous variable y_1 (economic vulnerability) as an explanatory variable.

The first specification with ProGres variables only does not require new data collection and uses variables already collected from refugees upon registration. The data are also updated on a regular basis by UNHCR. For this first model specification, the sets of explanatory variables are

$$x_1 = (HH\ size, homogeneous, Access\_Phone, HOH\ education, valuable\ assets, HHshares),$$
$$x_2 = (HH\ size, homogeneous, HOH\ education, HHshares),$$

where HHshares contains empirical proportions of household members related to age, gender, employment and disability status, and education level of the head of household (HOH). Variable *homogeneous* is introduced to capture the influence of the density of refugees from the same district of origin (in Syria) and living now in the same district (in Lebanon) as the considered household. It is computed as a proportion as well, using the information from ProGres on places of origin and current residence. To capture the role of having a means of communication, we include a dummy variable *Access_Phone*, equal to 1 if the household has positive expenditures on cell (GSM) or land-line phone(s). The possession by the household of appliances (or other household durable goods) with a significant market value (in case of sale on a local market) is captured by the variable *valuable assets*. As it is assumed to influence only economic vulnerability and not food insecurity, it does not include appliances that can be used for home cooking, and is not included in the list of variables x_2.

The second specification includes community-specific averages at the district level, including access to drinking water, sanitation status, crowdedness index, share of (as well as of heads of) households with chronic disease, share of households receiving medical care.

The list of variables and indicators included in the targeting models is included in the tables below. They include household (HH) characteristics, shares of members with certain characteristics, housing conditions, location characteristics and other socio-demographic indicators. Variables are either available in VASyR 2015 only or in both VASyR 2015 and ProGres. Community level indicators have also been calculated based on VASyR 2015 variables and are appended to the corresponding cases in ProGres. For example, the share of HH living in crowded conditions was calculated for each of the 26 districts using the VASyR 2015 data set. The calculated share was then appended to each case in ProGres by corresponding district. Even though the variables needed to calculate community level indicators are not available in ProGres, they can be calculated based on the VASyR 2015 data set or any newer nationally representative sample and appended to cases in ProGres.

Table 3 reports descriptive statistics for variables used in the various model specifications.

3 Results

3.1 Estimation results

We now present estimation results for our model specifications with increasing information used to represent simultaneously food insecurity and economic vulnerability: model (I) with ProGres variables only, and model (II) adding district-level indicators. For each model considered, random effects at district level are accounted for, as described above, to evaluate the log-likelihood function of the bivariate ordered Probit. Parameter estimates are presented in Table 4. To interpret correctly the sign of the effect of a given explanatory variable, It has to be remembered that a higher value of the dependent variable *percap_exp_quant* corresponds to a lower economic vulnerability, while a higher value of the variable *fsgradient* is associated with a higher degree of food insecurity.

In both model specifications, the exogeneity assumption for economic vulnerability in the equation of food insecurity is strongly rejected, with a *p*-value of the Wald test less than 0.001 in all cases. The level of economic welfare (as measured by the empirical quantile of household expenditures per head) is negative and significant in all specifications for food insecurity, which was expected. The null hypothesis of no-correlation between random terms in the system of equations (food insecurity and economic vulnerability), given covariates

Table 3 Descriptive statistics of the sample

Variable	Specification(s)	Mean	Standard Deviation
fsgradient	(I), (II), HH	4.9506	2.198
percap_exp_quant	(I), (II), HH	5.5049	2.8737
HHsize	(I), (II), HH	5.2099	2.3289
HHsize2	(I), (II), HH	32.7195	31.8355
homogeneous	(I), (II), HH	0.1026	0.0954
hoh_education_level = intermediate	(I), (II), HH	0.187	0.39
hoh_education_level = none	(I), (II), HH	0.1577	0.3645
hoh_education_level = primary_school	(I), (II), HH	0.4379	0.4962
hoh_education_level = read_write	(I), (II), HH	0.0927	0.2901
hoh_education_level = secondary_school	(I), (II), HH	0.0732	0.2606
hoh_education_level = technical	(I), (II), HH	0.0156	0.1239
hoh_education_level = university	(I), (II), HH	0.0358	0.1859
Access_Phone	(I), (II), HH	0.8963	0.3048
valuable_assets$^{(*)}$	(I), (II), HH	0.9203	0.2709
less_than_5_share	(I), (II), HH	0.1973	0.1891
btw_5_and_17_share	(I), (II), HH	0.2904	0.2374
btw_51_and_70_share	(I), (II), HH	0.0568	0.1521
aged_more_than_71_share	(I), (II), HH	0.0111	0.0701
btw_18_and_50_male_share	(I), (II), HH	0.2131	0.183
btw_18_and_51_female_share	(I), (II), HH	0.2311	0.1477
disabled_share	(I), (II), HH	0.0277	0.0908
water_access_hh_share	(II), Dist	0.8422	0.2226
drinkingwater_access_hh_share	(II), Dist	0.4659	0.2196
sanitation_access_hh_share	(II), Dist	0.4106	0.2226
crowded_hh_share	(II), Dist	0.5483	0.0927
chronichoh_hh_share	(II), Dist	0.2103	0.0672
chronic_hh_share	(II), Dist	0.4007	0.1003
receivehealth_hh_share	(II), Dist	0.1049	0.0562

3850 observations. $^{(*)}$: equation for economic vulnerability only. Specifications are as follows. (I): ProGres variables only; (II): (I) + district variables. HH and Dist: evaluated at household and district level respectively t4.1 **Table 4** Estimation results. Simultaneous structural equations, ordered Probit

	Dep. variable	Percap_exp_quant		fsgradient	
t4.3		(I)	(II)	(I)	(II)
t4.4	Percap_exp_quant	–	–	−0.0773***	−0.0916***
t4.5		–	–	(10.92)	(12.94)
t4.6	HHsize	−0.3322***	−0.3566***	−0.1134***	−0.1334***
t4.7		(11.87)	(12.69)	(4.14)	(4.87)
t4.8	HHsize2	0.0158***	0.0168***	0.0028	0.0036*
t4.9		(8.28)	(8.76)	(1.46)	(1.90)
t4.10	homogeneous	−0.2190	−0.5330***	−0.0636	−0.2383
t4.11		(1.22)	(2.93)	(0.34)	(1.26)
t4.12	hoh_edu1	−0.2769***	−0.3098***	0.1186	0.0934
t4.13		(2.76)	(3.07)	(1.22)	(0.96)
t4.14	hoh_edu2	−0.5001***	−0.5193***	0.3182***	0.3021***
t4.15		(4.90)	(5.07)	(3.22)	(3.05)
t4.16	hoh_edu3	−0.3359***	−0.3841***	0.1589*	0.1300
t4.17		(3.50)	(3.99)	(1.72)	(1.41)
t4.18	hoh_edu4	−0.3885***	−0.3866***	0.3264***	0.3193***
t4.19		(3.61)	(3.57)	(3.11)	(3.04)
t4.20	hoh_edu5	−0.2840**	−0.3600***	0.1496	0.1034
t4.21		(2.56)	(3.23)	(1.39)	(0.96)
t4.22	hoh_edu6	0.0694	−0.0093	0.0959	0.0443
t4.23		(0.42)	(0.06)	(0.60)	(0.28)
t4.24	hoh_edu7	REF	REF	REF	REF
t4.25	Access_Phone	0.6275***	0.5617***	–	–
t4.26		(10.49)	(9.35)	–	–
t4.27	valuable_assets	0.2831***	0.2789***	–	–
t4.28		(4.38)	(4.31)	–	–
t4.29	less_than_5_share	0.3898	0.0540	0.1424	−0.0120
t4.30	btw_5_and_17_share	(0.39)	(0.05)	(0.14)	(0.01)
t4.31	btw_51_and_70_share	0.6455	0.3642	0.5532	0.4298
t4.32	aged_more_than_71_share	(0.64)	(0.36)	(0.54)	(0.42)
t4.33	btw_18_and_50_male_share	1.4382	1.0899	0.3310	0.1903
t4.34	btw_18_and_51_female_share	(1.42)	(1.08)	(0.33)	(0.19)
t4.35	less_than_5_share	1.6759	1.3100	0.6336	0.4870
t4.36	btw_5_and_17_share	(1.62)	(1.26)	(0.61)	(0.47)
t4.37	btw_51_and_70_share	2.0983**	1.7359*	0.3456	0.1807
t4.38	aged_more_than_71_share	(2.08)	(1.72)	(0.34)	(0.18)
t4.39	btw_18_and_50_male_share	1.1717	0.8970	0.5462	0.4423
t4.40		(1.16)	(0.89)	(0.54)	(0.44)
t4.41	disabled_share	−0.3245*	−0.3471*	0.2731	0.2654
t4.42		(1.74)	(1.85)	(1.47)	(1.43)
t4.43	P	0.0664 (1.05)	0.2174*** (4.46)	LR = 1.07 (p-value = 0.31)	LR = 17.61 (*p*-value = 0.00)

3771 observations. Random effects for districts of origin and destination in all specifications. Specifications are as follows. (I): ProGres variables only; (II): (I) + district variables. See Appendix Table 12 for list of variables. *t*-statistics are in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. ρ is the correlation coefficient between both equations; LR is the Likelihood Ratio test (distributed as a χ^2 (1) under the null assumption of independence)

and under the random effect specification, is rejected at the 5% level of confidence for model specification (II) (with district variables) but is not rejected for model specification (I).

The household size is strongly significant and it has a negative effect in all model specifications for both equations, and it is decreasing and convex in specifications (I) and (II) for

Table 5 Targeting effectiveness, food security

Food Insecurity line = (fsgradient = 4)	Coverage of the food insecure (1)	Under-coverage (2)	Leakage (3)	Targeting differential = (1)–(3)
Currently assisted	74.23	25.77	38.98	35.25
ProGres variables	52.38	47.62	32.76	19.62
ProGres + district variables	55.81	44.19	32.40	23.41

Undercoverage is percent of poor individuals that do not receive transfer

Leakage is percent of individuals that receive transfer and are not poor

The targeting differential is the difference between the coverage rate and the participation rate for non-poor

economic vulnerability, while *HHsize2* is significant only at the 10% level for food insecurity and specification (II). All else being equal, a larger size of household decreases per head expenditures and it also increases food security. A possible explanation for such finding is that there are economies of scale in food consumption within the household (less food per head for larger households, resulting in a lower food insecurity index).

Let us now turn to the variable *homogeneous*, measuring the extent to which refugees of the same district of origin tend to regroup in the host country (Lebanon). This variable is significant only in the equation for economic vulnerability for the last specification (II) (with district variables), where it is negative, while it does not explain food insecurity. Such a finding indicates that a higher proportion of refugees from the same Syrian district and living now in the same Lebanese district has a negative effect on economic welfare. Hence, to cope with economic vulnerability, the strategy of refugees consisting in living close to households of the same origin is not improving their economic status (although it does not modify the food insecurity status). A possible interpretation of such result is a negative one: it may well be the heterogeneity of households in terms of origin (in Syria) that is profitable to refugees, instead of a greater concentration of individuals coming from the same geographical area.

Consider now education level as a determinant of food insecurity and economic vulnerability. The reference for education level is the higher category, i.e., university degree, and all other binary variables for education are to be interpreted with respect to this maximum education level. Estimation results show that a higher educational level (university level being used as reference in both equations) tends to decrease economic vulnerability (equivalently, to increase economic welfare), and it also has a negative impact on food insecurity. This is expected, as a higher educational level may be associated with a greater ability to cope with changing conditions of access to food as well as less economic vulnerability in general. The possession of at least one cell phone, captured by variable *Has_Cell_Phone*, has a positive and significant effect on economic welfare in both specifications. This can be interpreted by the fact that such portable communication device is an essential means for accessing information on economic opportunities (e.g., informal work), leading to less economic vulnerability.

The possession of valuable assets in the equation for economic vulnerability is positive and significant (it was omitted from the food insecurity equation to achieve identification, as was variable *Access_Phone*). This can illustrate the fact that households with valuable assets, that can be sold on formal of informal markets or between neighbours, are less vulnerable,

Table 6 Targeting effectiveness, economic vulnerability

Poverty line = (60% percentile of HH expenditure, around 114 USD)	Coverage of the poor (1)	Under-coverage (2)	Leakage (3)	Targeting differential = (1)–(3)
Currently assisted	81.20	18.80	25.21	55.99
ProGres variables	84.60	15.40	19.38	65.32
ProGres + district variables	85.66	14.34	18.75	66.91

Undercoverage is percent of poor individuals that do not receive transfer

Leakage is percent of individuals that receive transfer and are not poor

The targeting differential is the difference between the coverage rate and the participation rate for non-poor

Table 7 Targeting Accuracy, food security

Food Insecurity line = (fsgradient = 4)	Total	Food Insecurity Status		
		FI	FS	BPAC
Currently assisted	100	59.02	40.98	45.81
ProGres Variables	100	66.24	33.75	51.38
ProGres + district variables	100	66.47	33.52	54.68

Benefits' incidence is: (Sum of all transfers received by all individuals in the group)/(Sum of all transfers received by all individuals in the population). Aggregated transfer amounts are estimated using household size-weighted expansion factors

such assets being used as more or less liquid savings. Consider finally demographic characteristics of the household, represented by proportions or shares of household members. The only significant demographic variables (although at the 10 or 5% level only) in the equation of *percap_exp_quant* are the proportion of household members between 51 and 70 (*btw_51_and_70_share* with a positive effect) and the proportion of disabled household members (*disabled_share*, with a negative effect). Household shares are not significant in the equation for food insecurity, indicating the recursive nature of such demographic variables, which affect food insecurity only through economic vulnerability.

3.2 Targeting effectiveness and accuracy

The purpose of this section is to examine the performance of our procedure for improving the targeting of poor and food-insecure households. To do this, we need to evaluate the effectiveness and accuracy of policies using regional and district-level aggregates for targeting, compared with the actual policy of the World Food Program in Lebanon.

Consider first the relationship between food insecurity and poverty at the household level. This link is conceptually clear, especially in an urban context where economic access to food (purchasing power) is the dominant factor in food security. Hence, a particular dimension of food insecurity (alongside limited availability, low stability and insufficient utilization, see Dilley and Boudreau 2001) can be defined as "food poverty", that is, the status of a household regarding its access to food as a direct function of its purchasing power. Assuming a relationship with available income for access to food when defining food poverty, then it is expected that food poverty and overall poverty can be identified as different points on the same scale of income, and that people in need of food would be a smaller subset of those in overall poverty.

Table 8 Targeting Accuracy, economic vulnerability

Poverty line = (60% percentile of HH expenditure, around 114 USD)	Total	Poverty Status		
		P	NP	BPAC
Currently assisted	100	72.83	27.17	66.42
ProGres Variables	100	79.01	20.98	75.03
ProGres + district variables	100	80.16	19.83	75.75

In percent. P: poor; NP: non-poor; BPAC: Benefits' incidence is: (Sum of all transfers received by all individuals in the group)/(Sum of all transfers received by all individuals in the population). Aggregated transfer amounts are estimated using household size-weighted expansion factors

Let us first distinguish between individuals living either under or above some poverty line that identifies food poverty, and which can be defined by the Minimum Expenditure Basket (MEB)[2] set at 114$ per person per month which includes the cost of food plus other needs. However, since we are measuring food security and not food poverty with our newly developed index, we expect to find some discrepancy between food insecure and poor households. And because we are using ordered values for economic vulnerability to represent quantiles of household expenditures, we can only approximate such value. In fact, the poverty line discussed above corresponds roughly to the 60% quantile of the dependent variable, so that we consider all households below such quantile in our sample as poor. Note that the 60% quantile in consumption expenditure over the whole sample is different in general from the value corresponding to 60% of a given household's consumption expenditure.

As for food insecurity, we consider the fourth gradient value of the variable fsgradient as the threshold delimiting food insecurity. To explore the robustness of our method, we also tested an alternative food insecurity threshold, by considering the third value fsgradient =3 instead of 4. Results regarding food insecurity were very similar in terms of targeting accuracy.

Tables 5, 6, 7 and 8 present the performance of the targeting policy using only ProGres or using district-level aggregates as well, regarding food insecurity and economic vulnerability. Such performance is measured by the effectiveness and the accuracy of the targeting policy in both cases, which account for coverage, under-coverage and leakage rates of each policy (actual, ProGres variables, ProGres and district-level variables). A convenient indicator to measure accuracy of the targeting policy is the benefits' incidence. It is the transfer amount received by a group (in this case, poor or food insecure households) as a proportion of total transfers received by the population. Although the WFP Vulnerability Score leads to the highest percentage of transfers going to poor versus non-poor households, all two ProGres models have a higher Balanced Poverty Accuracy Criterion (BPAC), defined as Poverty

[2] The Minimum Expenditure Basket (MEB) is an index that is used to construct poverty lines in various contexts, including for refugee populations. It is emerging as a primary index to develop a cost and market based expression of minimum needs of refugees in any given country. It broadly follows the notion of a "cost of basic needs approach" as outlined in the World Bank Poverty Manual from 2005.

Accuracy minus the absolute difference between under-coverage and leakage. The model with ProGres and district-level variables for economic vulnerability has the best performance in terms of having the highest coverage, lowest leakage, highest percentage of transfers to the poor and highest BPAC.

The targeting effectiveness of the economic vulnerability models was much higher than for the food security models (the first model specification only includes variables collected during registration, while the second includes district level variables (geographic aggregates). The coverage rate of poor households reached 84.60% and 85.66% in the economic vulnerability model (see Table 6), compared with 52.38% and 55.81% in both model specifications of the food security model (see Table 5). Likewise, targeting accuracy was better in the economic vulnerability models, where the BPAC reached 75.03 and 75.75 compared with 51.38 and 54.68 in the food security model (see Tables 7 and 8).

4 Discussion

Using recent micro data from Syrian refugees in Lebanon, the paper investigated the empirical relationship between food insecurity and economic vulnerability at the household level. By estimating a system of structural equations for food insecurity and economic vulnerability, we showed how community-based variables such as population density and homogeneity of refugee households with respect to districts of origin and of arrival (residence) can improve the targeting effectiveness of aid programs, notably food aid. This is particularly important for increasing the performance of food aid policies when budgets are limited and/or decreasing.

A major result of interest to policy makers is that regional and community-based aggregates can be used to improve the targeting effectiveness of aid programs, e.g., food aid by the World Food Program dedicated to refugee population. Our results confirm that using such aggregates can augment the Balanced Poverty Accuracy Criterion, specially in our case, in terms of targeting effectiveness and accuracy for economically-vulnerable households.

With the cost of construction of aggregate indicators being in general less than individual-level data collection, a more accurate targeting of poor households may help attaining poverty alleviation objectives at a lower cost, when policy makers are faced with significant costs of poor households identification. As for poor and food-insecure households, being better targeted from the start allows them to benefit from a more efficient food aid system by, e.g., optimizing follow-up visits for in-depth monitoring.

By helping to reduce under-coverage and leakage of food and cash assistance programs., the empirical procedure considered in this paper can be used for policies based on in-kind as well as on cash transfers, because its purpose is to help identifying food-insecure and/or economically-vulnerable households, independently from the vector of aid.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix 1: Computation of the Food Insecurity Indicator

As part of the overall WFP vulnerability score, the food security sector score is constructed from three variables: food consumption score, food expenditure share and coping strategies index. The resulting score is converted into ordinal classes (categories) according to a formula developed by WFP VAM.[3] This score has been derived through an iterative process, and is based on several endogenous variables, which would be problematic in predictive models of food insecurity. We reviewed the VASyR dataset and considered all food security related variables in the dataset, to be used as potential food security outcomes (*y)* in a targeting formula.

The following indicators were considered, and were constructed according to standard WFP methods (World Food Program 2009):

– Food Consumption Score (FCS) *– a measure of quality of food utilization at household level; widely used to establish prevalence of food insecurity*
– Child Diet Diversity Score *– access to food quality by the most vulnerable*
– Coping Strategies Index (CSI) *– a measure of household economic access to food, food quality and food quantity; used in targeting food assistance in various contexts*
– Reduced Coping Strategies Index (rCSI)- *cross culturally validated measure of access to food*
– Food Expenditure Share *– a measure of household economic access to food*

It has been highlighted that reliance on a single measure which captures one dimension of food insecurity can misclassify the food insecure, and that combining indicators can improve the measurement of food insecurity (see Maxwell et al. 2013; Jones et al. 2013).

FCS and dietary diversity tend to capture elements of diet quality and diversity, whereas CSI and rCSI reflect quantity or sufficiency. Of these, child dietary diversity was not further explored as this would have reduced the sample of the dataset to households with children under the age of 2 years only. The Coping Strategies Index (CSI) asks a series of questions about

[3] WFP VAM Targeting verification criteria document

Table 9 Cross classification of bootstrapped quantiles of FCS and rCSI, % of households falling into each category

	rCSI Q1	**rCSI Q2**	**rCSI Q3**	**rCSI Q4**	**rCSI Q5**
FCS Q1	5.51	2.82	2.58	1.74	2.52
FCS Q2	4.98	2.80	3.24	2.53	3.44
FCS Q3	5.33	4.08	4.95	4.02	3.62
FCS Q4	5.59	3.40	4.28	5.46	3.90
FCS Q5	4.76	3.39	4.71	5.22	5.13

how households manage to cope with a shortfall in food for consumption and consists of a numerical score. It was not possible to construct the CSI according to standard methods as the VASyr 2015 posed the coping strategies questions in a way that does not allow the computation of the full index. The reduced Coping Strategies Index is a subset of the CSI that focuses on five food-related coping strategies and results in a cross-culturally validated tool to assess access to food. As the rCSI has been shown to reflect food insecurity as well as the full CSI, the rCSI was considered instead. Conceptually, we considered food expenditure as an economic determinant of food insecurity and therefore used it to validate the food security measure rather than as a component of the measure itself.

We therefore used FCS and rCSI as proxies of food quality and quantity, and used an empirical approach to derive cut-offs for relative vulnerability to food insecurity within this population, rather than international cut-offs developed for use in acute emergency settings.

Using FCS and rCSI as continuous variables, we derived both empirical and bootstrapped quantiles for each of the variables. As both of these approaches yielded similar results, we used the bootstrapped data in order not to impose restrictions on quantiles.

The simplest approach to combine the two variables was to cross classify these quantiles in the derivation of a food insecurity gradient, as has been done by others elsewhere (Maxwell et al. 2013). This cross classification yields a gradient of vulnerability to food insecurity. Considering rCSI Q1 to be the quantile with lowest coping, and FCS Q1 to be that with highest food consumption score, cases falling in the top left cell in Table 1 therefore have the lowest vulnerability to food insecurity. Conversely, cases falling in the bottom right cell (rCSI Q5 and FCS Q5) have the highest vulnerability to food insecurity.

In order not to impose arbitrary cut-off lines in classifying vulnerability to food insecurity, we tested the food insecurity gradient against economic variables conceptualized as determinants of vulnerability to food insecurity; food expenditure, total expenditure, extreme poverty (below SMEB) & overall poverty (below MEB).

Assuming a food insecurity gradient across quantiles of rCSI and FCS, leads to 9 levels of vulnerability to food insecurity (along diagonals of Table 1). Table 2 displays average food and total monthly household expenditures (in USD), proportion of households categorized as poor and extreme poor by food insecurity gradient in the sample. Data show that, as the food insecurity gradient increases, mean monthly food expenditures and total expenditures decrease, while poverty, extreme poverty and percentage share of food expenditure increase.

Table 10 Economic characteristics of households at different levels of the food insecurity gradient

Food security gradient	Percent of population	Monthly food expenditures (mean, USD)	p-value*	Total monthly expenditures (mean, USD)	p-value*	Below poverty line (%)	Below extreme poverty (%)
1	5.51	246.10		651.80		37.6	2.2
2	7.8	236.65	0.55	574.80	0.06	51.3	2.9
3	10.71	203.40	0.00	500.95	0.03	53.2	8.1
4	14.65	177.15	0.01	423.69	0.00	66.5	8.0
5	18.16	155.34	0.01	417.43	0.01	73.8	10.2
6	15.13	147.92	0.34	347.54	0.12	78.3	16.3
7	13.79	133.55	0.07	297.07	0.01	86.0	17.8
8	9.12	128.60	0.54	295.18	0.92	87.2	19.5
9	5.13	136.19	0.45	310.46	0.51	85.1	12.5

*p-values for differences between means across gradients

Table 11 Cross classification of bootstrapped quantiles of FCS and rCSI, according to gradient thresholds

	rCSI Q1	**rCSI Q2**	**rCSI Q3**	**rCSI Q4**	**rCSI Q5**
FCS Q1	5.51	2.82	2.58	1.74	2.52
FCS Q2	4.98	2.80	3.24	2.53	3.44
FCS Q3	5.33	4.08	4.95	4.02	3.62
FCS Q4	5.59	3.40	4.28	5.46	3.90
FCS Q5	4.76	3.39	4.71	5.22	5.13

Based on an analysis of differences in mean monthly food and total expenditures across gradients, thresholds of vulnerability to food insecurity were derived. In brief, where there were significant differences in expenditures across gradients, a threshold line was drawn, yielding five categories of vulnerability to food insecurity. Table 3 displays the cross classification of the bootstrapped quantiles, with thresholds drawn between gradients 2 and 3, 3 and 4, 4 and 5, and 6 and 7.

Appendix 2

Table 12 Description of variables

Variable	Description
fsgradient	Food Security gradient (ordered 1–9)
Percap_exp_quant	Household expenditure quantile (1–10)
HHsize	Number of household members
HHsize2	HHsize squared
Homogeneous	Proportion of sample HH in Lebanese district from same district of origin in Syria, see text
hoh_education_level = intermediate	1 if head of HH education level: intermediate
hoh_education_level = none	1 if head of HH education level: none
hoh_education_level = primary_school	1 if head of HH education level: primary
hoh_education_level = read_write	1 if head of HH education level: read and write
hoh_education_level = secondary_school	1 if head of HH education level: secondary
hoh_education_level = technical	1 if head of HH education level: technical
hoh_education_level = university	1 if head of HH education level: higher
Access_Phone	1 if HH expenditures on phone(s) are positive
valuable_assets	1 if HH possesses valuable assets (durable goods)
less_than_5_share	Proportion of HH members under 5 years of age
btw_5_and_17_share	Proportion of HH members aged 5 to 17
btw_51_and_70_share	Proportion of HH members aged 51 to 70
aged_more_than_71_share	Proportion of HH members aged >70
btw_18_and_50_male_share	Proportion of male HH members between 18 and 50
btw_18_and_51_female_share	Proportion of female HH members between 18 and 50
disabled_share	Proportion of disabled HH members
water_access_hh_share	Share of HH with access to sufficient amount of water for drinking, cooking, washing and toilet purposes
drinkingwater_access_hh_share	Share of HH with access to safe drinking water
sanitation_access_hh_share	Share of HH with access to flush toilets
crowded_hh_share	Share of HH living in crowded conditions
chronichoh_hh_share	Share of HH where Head of HH is chronically ill
chronic_hh_share	Share of HH who have one or more chronically ill members
receivehealth_hh_share	Share of HH who receive health care/drugs regularly

References

Agostini, C. A., & Brown, P. H. (2011). Cash transfers and poverty reduction in Chile. *Journal of Regional Science, 51*, 604–625.

Amarasinghe, U., Samad, M., & Anputhas, M. (2005). Spatial clustering of rural poverty and food insecurity in Sri Lanka. *Food Policy, 30*, 493–509.

Barrett, C. B. (2002). Food security and food assistance programs. In *Handbook of Agricultural Economics*, vol. 2, Chapter 40. Dordrecht: Elsevier Science.

Benfield, W. A. (2007). Optimum targeting: The Jamaica food stamp programme. *Social and Economic Studies, 56*(4), 121–164.

Besley, T., & Kanbur, R. (1993). The principles of targeting. In M. Lipton & J. Van der Gaag (Eds.), *Including the Poor*. Washington DC: World Bank.

Besley, T., & Kanbur, R. (1988). Food subsidies and poverty alleviation. *Economic Journal, 98*, 701–719.

Borjas, G. J. (2004). Food insecurity and public assistance. *Journal of Public Economics, 88*, 1421–1443.

Chung, K., Haddad, L., Ramakrishna, J., & Riely, F. (1997). *Identifying the food insecure: The application of mixed-method approaches in India*. Washington DC: International Food Policy Research Institute.

Clay, D. C., Molla, D., & Habtewold, D. (1999). Food aid targeting in Ethiopia: A study of who needs it and who gets it. *Food Policy, 24*(4), 391–409.

Dercon, S. (2006). Vulnerability: A micro perspective. In *Securing Development in an Unstable World* (François bourguignon, Boris Pleskovic and Jacques van der Gaag, eds.), pp. 117-146. Proceedings of the annual World Bank conference on development economics—Europe 2006.

Dilley, M., & Boudreau, T. E. (2001). Coming to terms with vulnerability: A critique of the food security definition. *Food Policy, 26*(3), 229–247.

Drummond, J., R. Khoury, S. Bailey, N. Crawford, L. Fan, R. Milhem, and S. A. Zyck (2015). *An Evaluation of WFP's Regional Response to the Syrian Crisis*, 2011-2014. Evaluation report OEV/2014/19, Overseas Development Institute for the WFP Office of Evaluation. Rome: United Nations.

Elbers, C., Fujii, T., Lanjouw, T., Özler, B., & Yin, W. (2007). Poverty alleviation through geographic targeting: How much does disaggregation help? *Journal of Development Economics, 83*(1), 198–213.

Glewwe, P. (1992). Targeting assistance to the poor: Efficient allocation of transfers when household income is not observed. *Journal of Development Economics, 38*(2), 297–321.

Haddad, L., Hoddinott, J., & Alderman, H. (Eds.). (1997). *Intrahousehold resource allocation in developing countries: Models, methods and policy*. Baltimore: Johns Hopkins University Press.

Hyman, G., Larrea, C., & Farrow, A. (2005). Methods, results and policy implications of poverty and food security mapping assessments. *Food Policy, 30*, 453–460.

IRIS Center. (2005). Note on assessment and improvement of tool accuracy. http://www.povertytools.org/training_documents/Introduction%20 to%20PA/Accuracy_Note.Pdf.

Jaynes, T. S., Strauss, H., & Yamano, T. (2001). Giving to the poor? Targeting of food aid in rural Ethiopia. *World Development, 29*(5), 887–910.

Jones, A. D., Ngure, F. M., Pelto, G., & Young, S. L. (2013). What are we assessing when we measure food security? A compendium and review of current metrics. *Advances in Nutrition: An International Review Journal, 4*(5), 481–505.

Kam, S. P., Hossain, M., Bose, M. L., & Villano, L. S. (2005). Spatial patterns of rural poverty and their relationship with welfare-influencing factors in Bangladesh. *Food Policy, 30*, 551–567.

Lipton, M., & Ravallion, M. (1995). Poverty and policy. In: Behrman J. & Srinivasan, T.N. (eds.), *Handbook of Development Economics*, vol. 3B. Amsterdam: Elsevier.

Maxwell, D., Coates, J., & Vaitla, B. (2013). *How do different indicators of household food security compare? Empirical evidence from Tigray*. Medford: Feinstein International Center, Tufts University.

Miller, C. L., Tsoka, M., & Reychert, K. (2011). The impact of the social cash transfer scheme on food security in Malawi. *Food Policy, 36*, 230–238.

Minot, N., & Baulch, B. (2005). Spatial patterns of poverty in Vietnam and their implications for policy. *Food Policy, 30*, 461–475.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2005). Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Journal of Econometrics, 128*, 301–323.

Sajaia, Z. (2007). *Maximum likelihood estimation of a bivariate ordered Probit model: Implementation and Monte Carlo simulations*. Washington, DC: World Bank.

Sen, A. (1995). The political economy of targeting. In D. van de Walle & K. Need (Eds.), *Public spending and the poor: Theory and evidence*. Maryland: The Johns Hopkins University Press.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Boca Raton: Chapman & Hall/CRC.

UNHCR, UNICEF & WFP, (2015). Vulnerability Assessment of Syrian Refugees in Lebanon.

Wilde, P., & Nord, M. (2005). The effect of food stamps on food security: A panel data approach. *Review of Agricultural Economics, 27*(3), 425–432.

World Food Program. (2009). *Comprehensive Food Security & Vulnerability Analysis Guidelines*. Rome: United Nations.

World Food Program. (2016). *Lebanon - Vulnerability Assessment of Syrian Refugees in Lebanon (December)*. Rome: United Nations.

Dr. Jad Chaaban graduated with a Ph.D. in economics from the University of Toulouse. He is now an Associate Professor of Economics at the American University of Beirut (AUB). His primary research interests are in development economics and industrial organization, and his extended interests include public economics of health, education and labor policies, in addition to agricultural and environmental economics. His work focuses mostly on using tools in micro and applied economics to tackle public policy issues that affect the lives of the most vulnerable groups in society.

Dr. Hala Ghattas holds a Ph.D. from the University of London and is currently Assistant Research Professor and Associate Director of the Center for Research on Population and Health at the American University of Beirut. Her research interests include food security, nutrition and health status of refugees and marginalized populations.

Alexandra Irani holds a MSc in Agricultural Economics and Development from the American University of Beirut, Lebanon. She is currently working on several projects including socio-economic status, food insecurity and poverty of refugees in Lebanon (from Palestine and Syria), and she conducted research for the 2015 Arab Human Development Report.

Alban Thomas is a Senior Researcher and the head of the Social Science Division at INRA (French Institute for Agricultural Research). An environmental and agricultural economist, his research topics include the empirical analysis of environmental impacts from agriculture, the evaluation of environmental policies, water use and tariff design, and the adoption of innovative agricultural systems. He has co-ordinated several scientific expertises on water and the environment for ministries and international institutions, and he is the co-ordinator of the interdisciplinary Inra-Cirad ten-year programme GloFoodS on transitions towards global food security.