VOICES
of the
HUNGRY

Measuring Food Insecurity through people's experiences

State-of-the-art and results from four pilot studies in Sub Saharan Africa

Carlo Cafiero and the VoH Team, FAO – ESS

Special thanks to Mark Nord and to Sara Viviani

Department for International Development

ROYAUME DE BELGIQUE
Affaires étrangères,
Commerce extérieur et
Coopération au Développement

ESS

Outline

I. Introducing a proper theory of measurement in the social sciences
- Objective (invariant) measurement of latent traits
- The Rasch model, as the cornerstone of Item-Response Theory (IRT)

II. Application to the severity of food insecurity
- Defining the concept of 'severity of food security' through the FIES
- Analyzing the responses to the FIES

III. Preliminary results of Voices of the Hungry pilot tests in Angola, Ethiopia, Malawi and Niger
- Validation of the tools
- Analysis of relationship between threshold levels and prevalence rates

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

A theory of measurement for social science

- In social sciences, the *objects* being measured are often rather elusive (ex.: “democracy”, “wellbeing”, “power”, “intelligence”, “food security”, etc.)
  - These have been termed “**latent traits**” or “**constructs**”
  - Existence is revealed only indirectly, through observable phenomena (behaviors, experiences, consequences, etc.)
- It is impossible to validate measures by reference to a “gold standard”
  - Validity **of the underlying concept**
    - By consensus on its usefulness in informing discussions
  - Validity **of the measurement tool**
    - Based on its theoretical foundation
  - Validity **of particular measures**
    - Performance of the tool in specific conditions

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

A theory of measurement for social science

- The validity of the measurement poses special questions as regarding the *objectivity* (invariance) of the measures
  - Measures obtained with one tool ought to be independent of the object being measured
  - Measures of the same object obtained with different tools, should be independent of the tool being used
- Advocates of the **Rasch model** claim that it is the *only* psychometrics model that fulfills these requirements
  - References: Rasch (1960), Fischer and Molenaar (2006), Engelhard (2013)

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

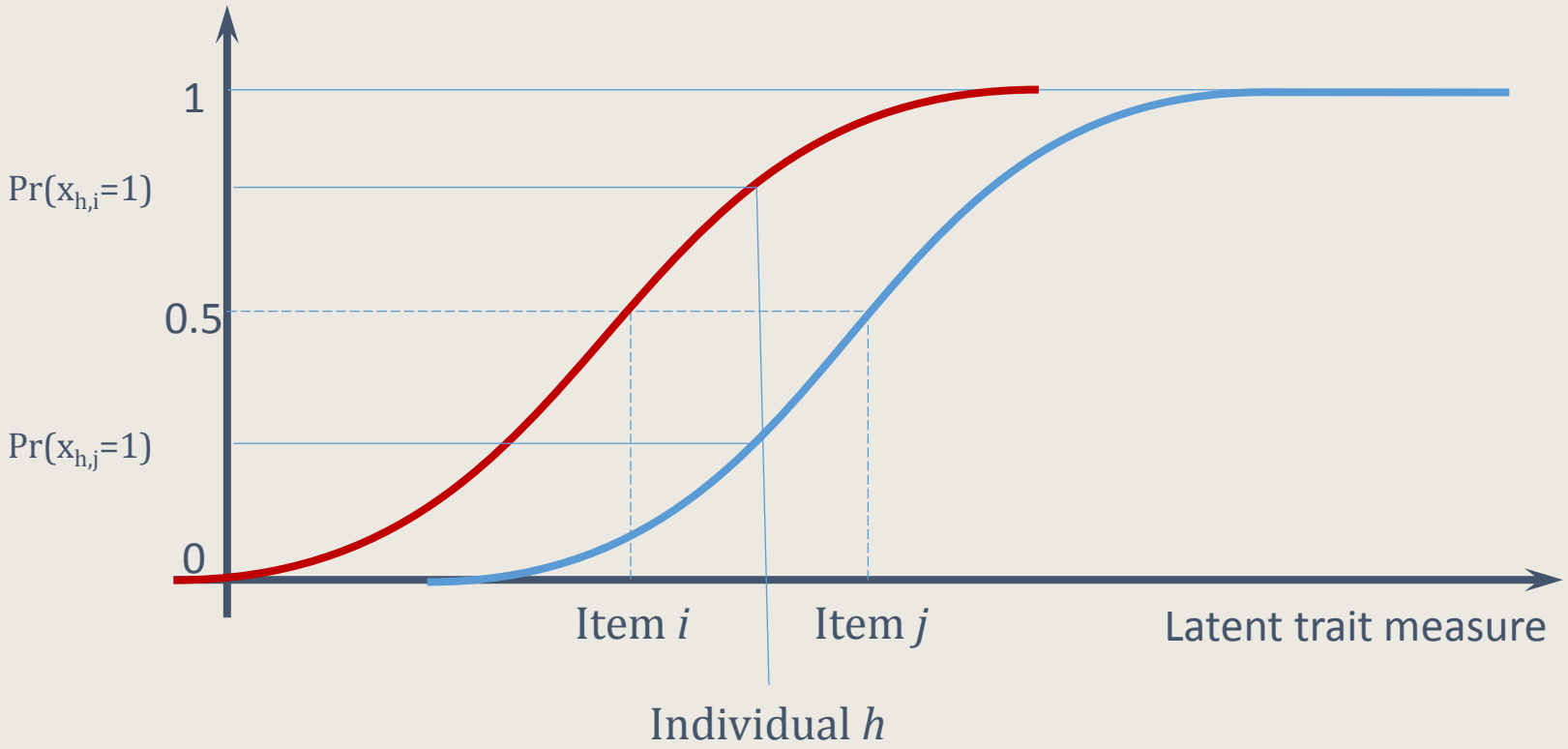
The Rasch model

- Developed to measure latent traits on individuals
  - Developed in educational testing and psychometrics (Rasch, 1960)
  - Known in econometrics as the *logit* model
- The probability of one individual affirming an item (or of making a decision, or of showing a symptom, etc.) depends on the distance between the item and the individual on the latent trait scale

$$\mathrm{Prob}\left(x_{h,i} = 1 \middle| h, i\right) = \frac{e^{h-i}}{1 + e^{h-i}}$$

- As it depends on the *distance* (not the absolute value of the measures) setting the "zero" is arbitrary
- Any affine transformation of the scale would yield the same ranking (arbitrariness of the scale for ranking purposes)
- With *n* items, only *n-1* item parameters can be estimated

The Rasch model



The Rasch model

- A special case of more general latent variable models

$$\mathrm{Prob}(x_i = 1|\boldsymbol{\theta}) = \phi(\boldsymbol{\theta})$$

- Extensions are available for *polytomous* models, where the observed variable is multinomial
- Framing it in probabilistic terms…
  - … recognizes that observations may be affected by errors
  - … provides the basis for *estimating* the parameters of the model by maximizing the likelihood of any observed set of data
    - For dichotomous data, i.e. $\{x_i\} \in \{0,1\}\ \forall\ i$, and assuming independence of the responses

$$L(X|\boldsymbol{\theta}) = \prod_i \phi(\boldsymbol{\theta})^{x_i} \times (1 - \phi(\boldsymbol{\theta}))^{1-x_i}$$

  - … allows for the analysis of goodness of fit as a way to tackle the **empirical validity** question in a rigorous way

The analytics of the Rasch model

- Quite simple in practice, yet sophisticated enough to address all the relevant questions of empirical validity
- Heuristically, it can be described as taking two passages through the data
  - First, by looking at the items, to determine their relative position on the scale
  - Then, by looking at the individuals, to determine their position on the scale

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

The Rasch model

Individuals	Items: I	II	II	…	N	Sum (raw score)
1	1	1	0	…	0	T^1
2	1	1	1	…	0	T^2
3	1	1	0	…	1	T^3
…	…	…	…	…	…	…
m	1	0	0	…	0	T^m
Sum	T^I	T^{II}	T^{III}		T^N	T

The logic underpinning the Rasch model

- It imposes the necessary restrictions to allow consistent estimates of the item parameters:
  - Conditionally on items' severity, only response patterns of the type {1, …, 1, 0, …, 0} are admissible.
  - Only non-extreme responses (i.e., at least one 0 and at least one 1) are informative
  - This implies that the raw scores (i.e., the sums of the elements in the response vector) is a *sufficient statistics* for the respondent severity
  - Under the truth of the model, and given the items' severity, each pattern of responses has a definite expected probability to be observed
  - Measures of *fit* can be produced by analyzing the expected probability of observed patterns, and by computing and analyzing "residuals" as the difference between the actual response and the expected probability of a positive response

Tests of the Rasch model

- Fit statistics
  - 0.8 – 1.2 : adequate fit
  - 0.7 – 1.3 : productive for measurement, though revealing potential problems
- Robustness of the severity items to split-samples
  - Invariance of the measure implies that the severity associated with the items should not depend on the particular set of respondents
- Factor analysis of the “residuals”
  - To detect the presence of additional dimensions not captured by the scale
- In case the Rasch model does not fit the data
  - Revise the choice of the items (excluding the problematic ones)
  - Use more sophisticated models to score cases
    - Probabilistic assignment of cases to classes
    - 2PLM, allowing for unequal discrimination among items

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Scaling and equating

- As the Rasch measure is scale invariant, and because different application may produce measures on different scales, comparison across samples requires bringing the measure to a common scale
  - A common scale can be developed, by equalizing the mean and standard deviation of a group of common items (at least 2, and up to all items in the scale)
  - If not all items are common items, raw scores cannot be compared across applications

Important: all items in the scale contribute to the measure.
It makes little sense to analyze each question separately.

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

The severity of food insecurity

The FIES:
a set of questions revealing experiences spanning the range of severity

	Food security	Food insecurity			
		mild	moderate		severe
experiences		**Worries**	Compromising food **quality and variety**	Compromising **food quantity**	**Hunger**
consequences	**Wellbeing**	**Welfare reduction** (Psychological costs, reduction of other essential expenses)	**Malnutrition** (obesity, micronutrient deficiencies, reduced work capacity)	**Undernutrition** (stunting, wasting)	**Starvation**

Department for International Development

ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement

ESS

back

The (current version of the) FIES

*"During the last 12 months, was there a time when, **because of lack of money or other resources**:*

1. *You were **worried** you would not have enough food to eat?*
2. *You were unable to eat **healthy and nutritious** food?*
3. *You ate only a **few kinds** of foods?*
4. *You had to **skip a meal**?*
5. *You **ate less** than you thought you should?*
6. *Your household **ran out** of food?*
7. *You were **hungry** but did not eat?*
8. *You went **without eating for a whole day?"***

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Linguistic and cultural adaptation

- Each question aims at detecting “typical” experiences associated with increasing severity
- The *meaning* of the question must be clearly understood by the respondent to give informative answers
  - Need to express it in the language of the respondent, and to adapt it the local cultural context
  - The importance of the “lack of resources” qualifier
  - And of proper training of the interviewers
- Despite all efforts in cultural and linguistic adaptation, estimated severity for some items may differ across countries
  - Equating must be based on a restricted number of common items
  - Probabilistic assignment of cases to food security classes to calculate prevalence levels, taking into consideration both common and idiosyncratic items

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Applying the FIES through the GWP

- Individual experiences
  - Consistent with the GWP reference population
  - Increases the reliability of the responses
  - Allows for analysis of possible gender discrimination
- 12 months reference period
  - Consistent with the need of an annual monitoring on a global scale
  - To ensure comparability of assessments across countries
- No follow up questions related to the frequency of experience
  - While undoubtedly important, frequency of occurrence is a *distinct dimension* of the food insecurity experience, calling for 2-dimensional classification

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Results from the four pilot studies

- The FIES was adapted to local languages in Ethiopia, Angola, Malawi and Niger
  - Reports of the adaptation are available on the VoH webpage
- Included in the 2013 round of the Gallup World Poll™
  - Nationally representative samples of the adult (15+) population in each country
  - Sample size: 1000
  - Face-to-face interviews in local languages
- Results have been analyzed to:
  - validate the concept by analyzing the fit of the Rasch model
  - determine the most appropriate procedure for scoring and classification
  - analyze robustness of prevalence estimates to threshold settings

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Results from the pilot studies: Angola

Item	Severity	Standard Errors	Infit	Outfit
WORRIED	-1.59	0.14	1.18	1.46
HEALTHY	-0.76	0.12	0.86	1.04
FEWFOOD	-0.37	0.11	1.07	1.28
SKIPPED	-0.24	0.11	0.87	0.74
ATELESS	-0.44	0.11	0.92	0.81
RUNOUT	0.25	0.11	0.86	0.72
HUNGRY	0.87	0.11	1.00	0.98
WHLDAY	2.27	0.13	1.18	1.63
Mean	0			
St. Dev.	1.09			
N (complete, non extreme)	505			

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Results from the pilot studies: Ethiopia

Item	Severity	Standard Errors	Infit	Outfit
WORRIED	-0.75	0.11	1.4	2.25
HEALTHY	-2.17	0.13	0.92	2.19
FEWFOOD	-2.63	0.14	0.96	1.11
SKIPPED	-0.02	0.11	0.74	0.71
ATELESS	-0.6	0.11	0.84	0.85
RUNOUT	1.36	0.12	0.76	0.69
HUNGRY	1.64	0.13	0.84	0.8
WHLDAY	3.16	0.17	1.1	3.7
Mean	0			
St. Dev.	1.84			
N (complete, non extreme)	597			

Department for International Development
ROYAUME DE BELGIQUE Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Results from the pilot studies: Malawi

Item	Severity	Standard Errors	Infit	Outfit
WORRIED	-0.49	0.13	1.11	1.17
HEALTHY	-0.36	0.13	1.16	1.36
FEWFOOD	-1.00	0.14	0.84	0.62
SKIPPED	0.22	0.12	0.97	0.88
ATELESS	-0.70	0.14	0.94	0.97
RUNOUT	0.12	0.12	0.9	0.81
HUNGRY	0.54	0.12	1.03	1.05
WHLDAY	1.67	0.12	1.04	1.14
Mean	0			
St. Dev.	0.79			
N (complete, non extreme)	423			

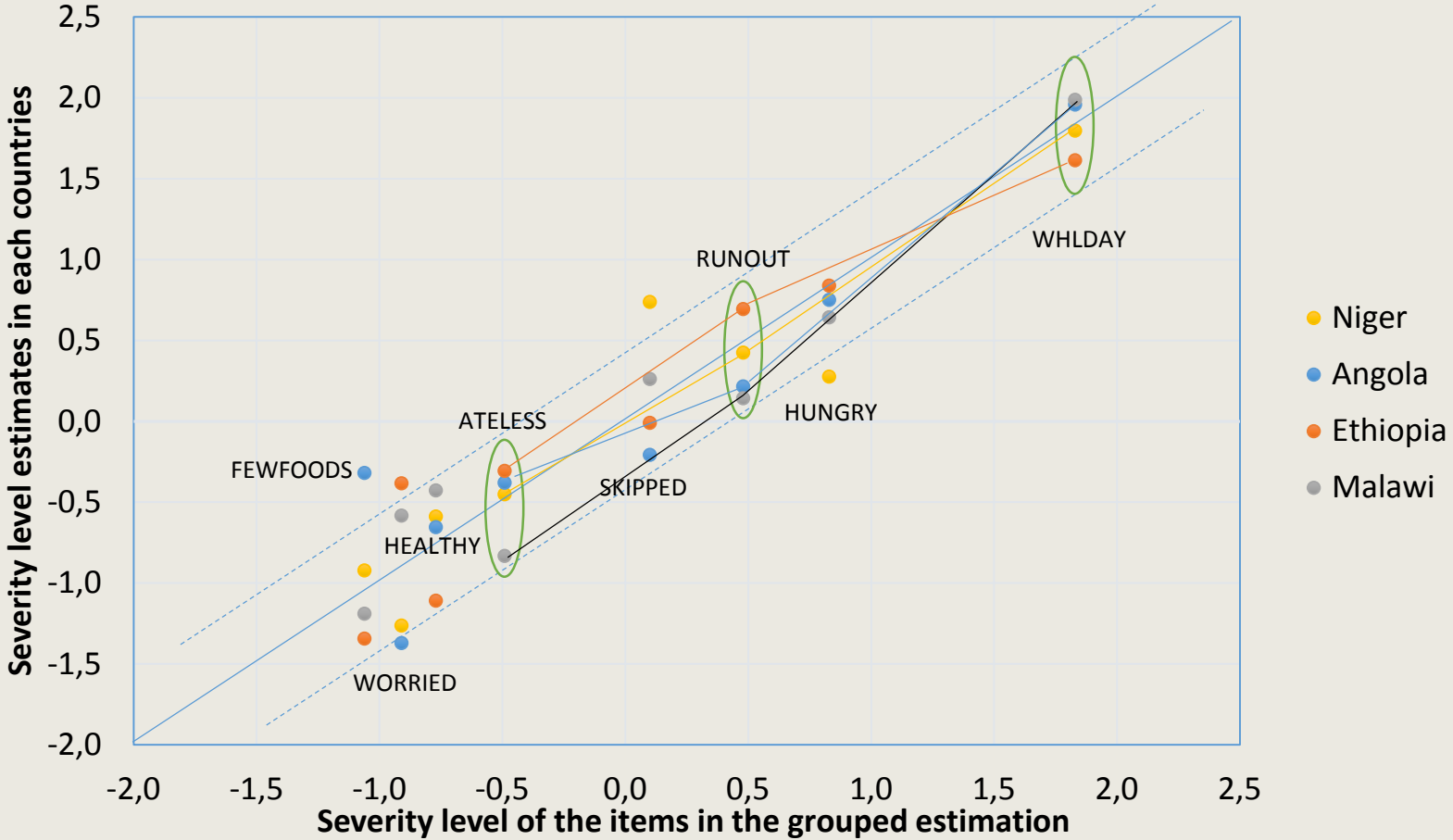
Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Results from the pilot studies: Niger

Item	Severity	Standard Errors	Infit	Outfit
WORRIED	-1.37	0.11	1.24	1.62
HEALTHY	-0.64	0.1	0.92	0.76
FEWFOOD	-1.00	0.11	1.01	0.9
SKIPPED	0.80	0.09	1.06	1.05
ATELESS	-0.49	0.1	0.98	1.02
RUNOUT	0.46	0.09	0.79	0.68
HUNGRY	0.30	0.09	0.82	0.75
WHLDAY	1.95	0.1	1.21	1.33
Mean	0			
St. Dev.	1.02			
N (complete, non extreme)	734			

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Comparison after rescaling, to identify common items



Department for International Development

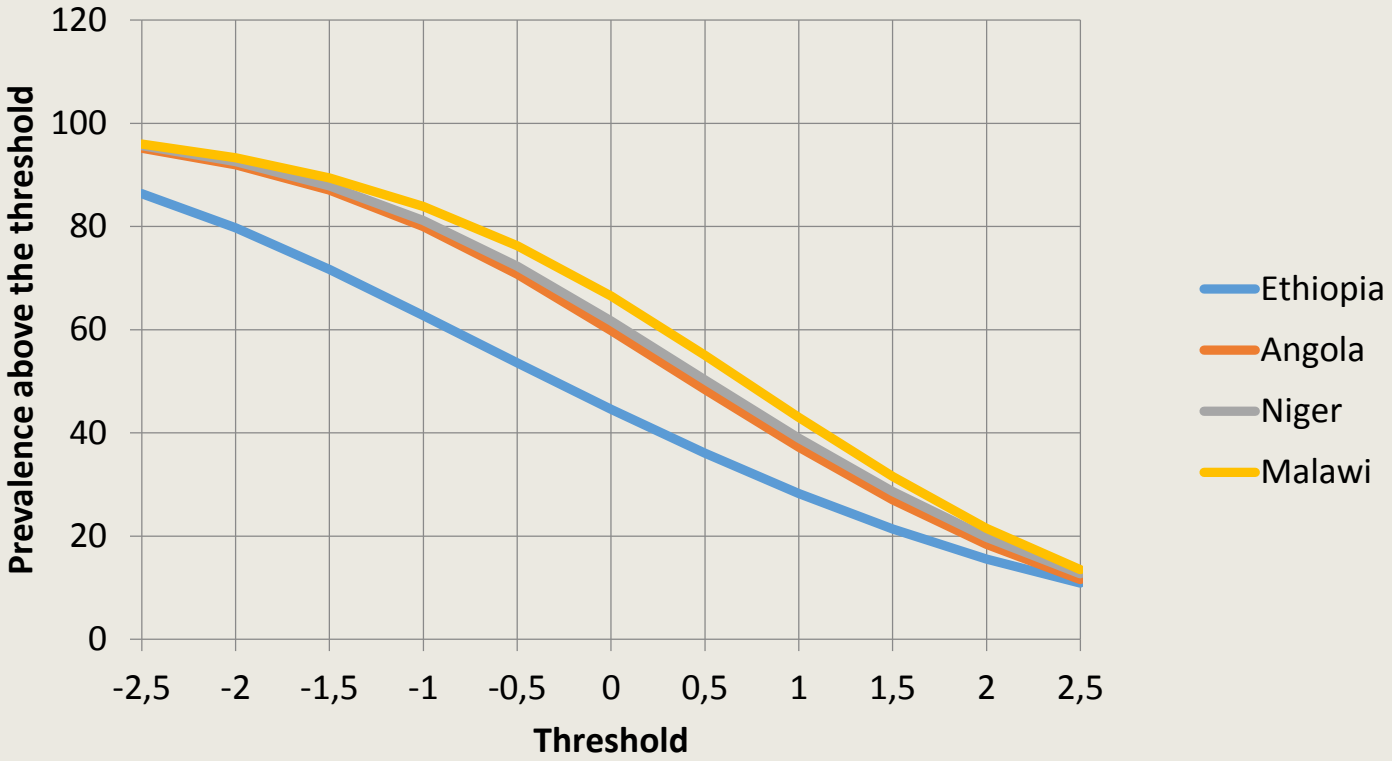
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement

ESS

Prevalence rates at different thresholds

	Thresholds										SD (within countries)
	-2	-1.5	-1	-0.5	0	0.5	1	1.5	2	2.5	
Grouped sample											
Prevalence (%)	**90.8**	**85.2**	**77.7**	**68.3**	**57.6**	**46.5**	**35.8**	**26.2**	**18.0**	**11.6**	
Index	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*	
Individual countries											
Ethiopia	**79.7**	**71.7**	**62.8**	**53.6**	**44.6**	**36.1**	**28.3**	**21.4**	**15.6**	**10.9**	
	88	*84*	*81*	*78*	*77*	*78*	*79*	*82*	*86*	*93.6*	5.6
Angola	**92.0**	**87.1**	**80.0**	**70.7**	**59.9**	**48.4**	**37.3**	**27.1**	**18.4**	**11.6**	
	101	*102*	*103*	*104*	*104*	*104*	*104*	*103*	*102*	*100*	1.4
Niger	**92.6**	**87.9**	**81.2**	**72.3**	**61.7**	**50.3**	**39.0**	**28.6**	**19.8**	**12.7**	
	102	*103*	*104*	*106*	*107*	*108*	*109*	*109*	*110*	*109*	3.1
Malawi	**93.3**	**89.4**	**83.9**	**76.3**	**66.5**	**55.1**	**43.0**	**31.5**	**21.5**	**13.5**	
	103	*105*	*108*	*112*	*116*	*118*	*120*	*120*	*119*	*116*	7.2
SD (between countries)	7.1	9.7	12.4	14.7	16.4	17.4	17.4	16.3	13.9	9.9	

Prevalence rates at different thresholds

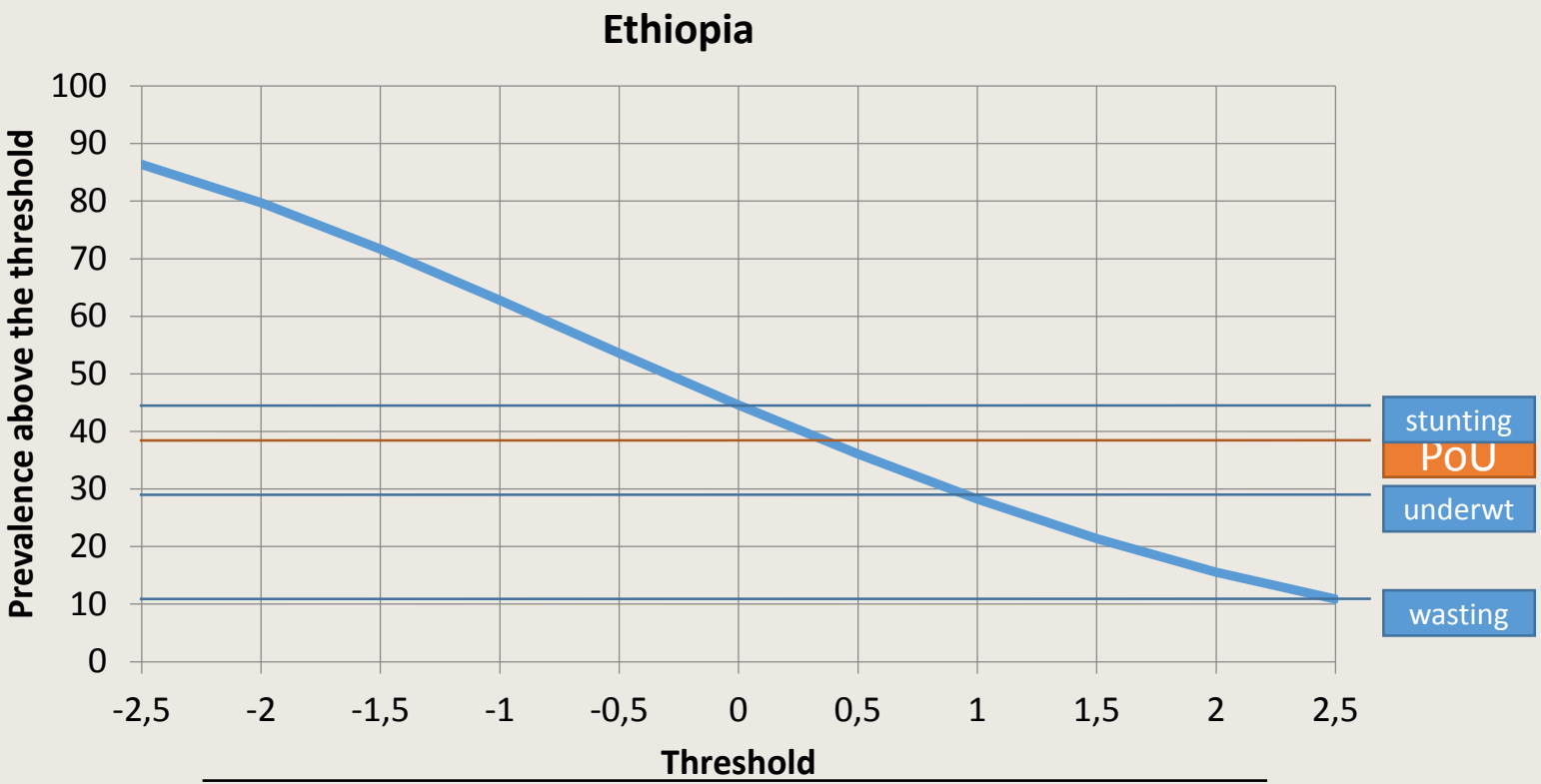


Department for International Development

ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement

ESS

Prevalence rates at different thresholds



	PoU (2011-13)	Child anthropometrics (% among under 5)			
		Wasting	Underweight	Stunting	(2011)
Ethiopia	37.1	10.1	29.2	44.2	

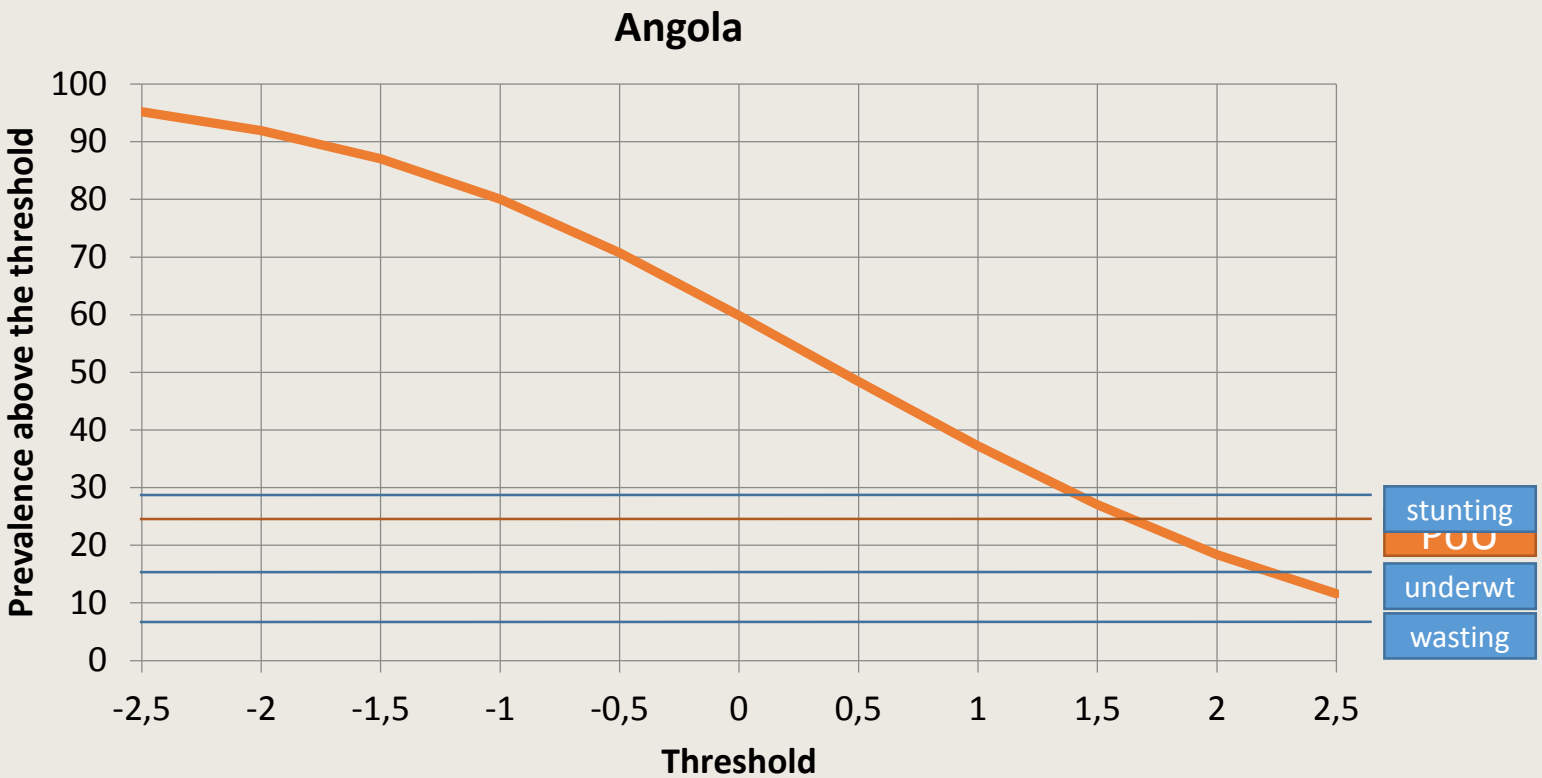
To the scale

Department for International Development

ROYAUME DE BELGIQUE Affaires étrangères, Commerce extérieur et Coopération au Développement

ESS

Prevalence rates at different thresholds



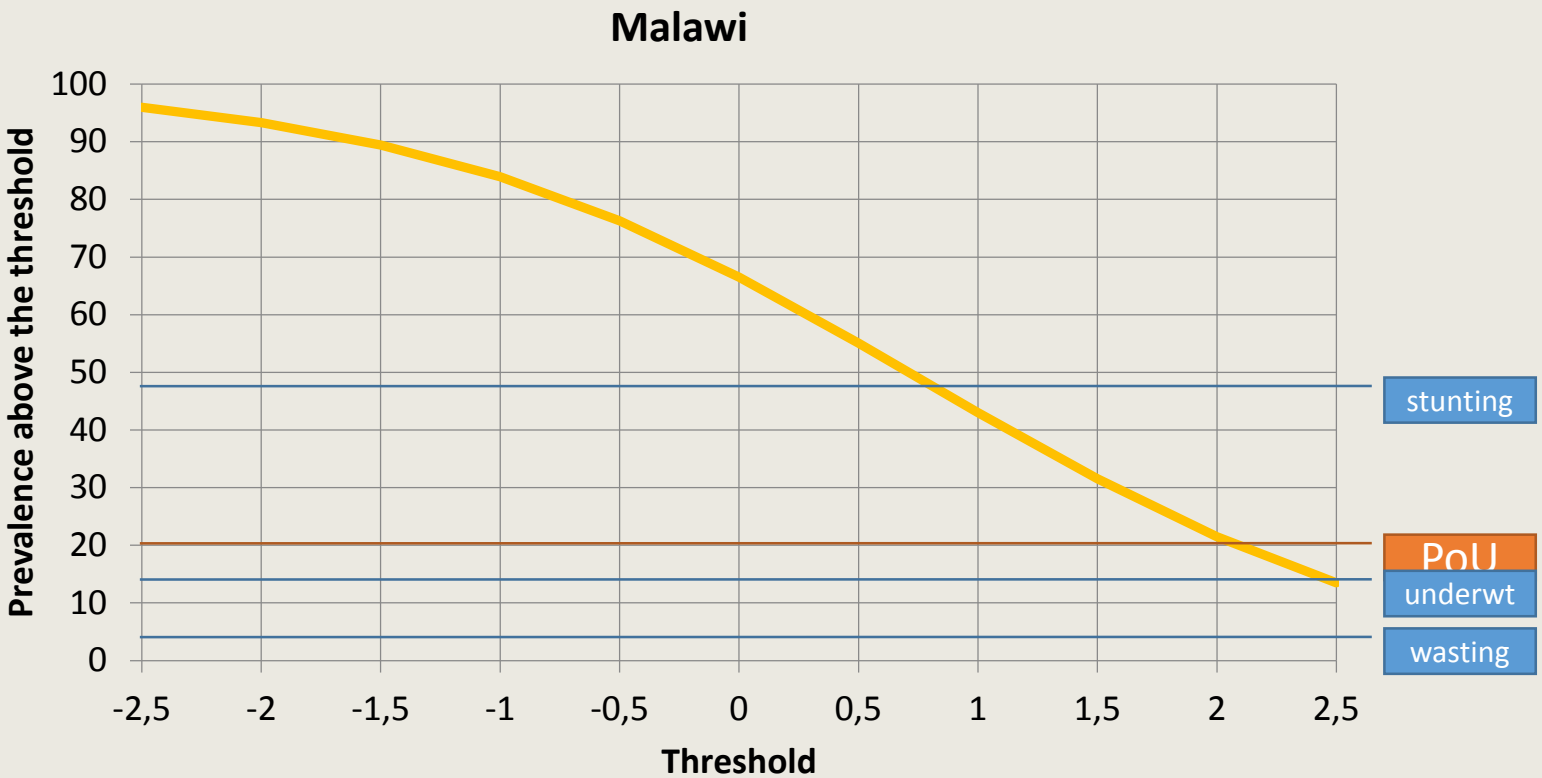
	PoU (2011-13)	Child anthropometrics (% among under 5)			
		Wasting	Underweight	Stunting	(2006)
Angola	24.4	8.2	15.6	29.2	

Department for International Development

ROYAUME DE BELGIQUE Affaires étrangères, Commerce extérieur et Coopération au Développement

ESS

Prevalence rates at different thresholds



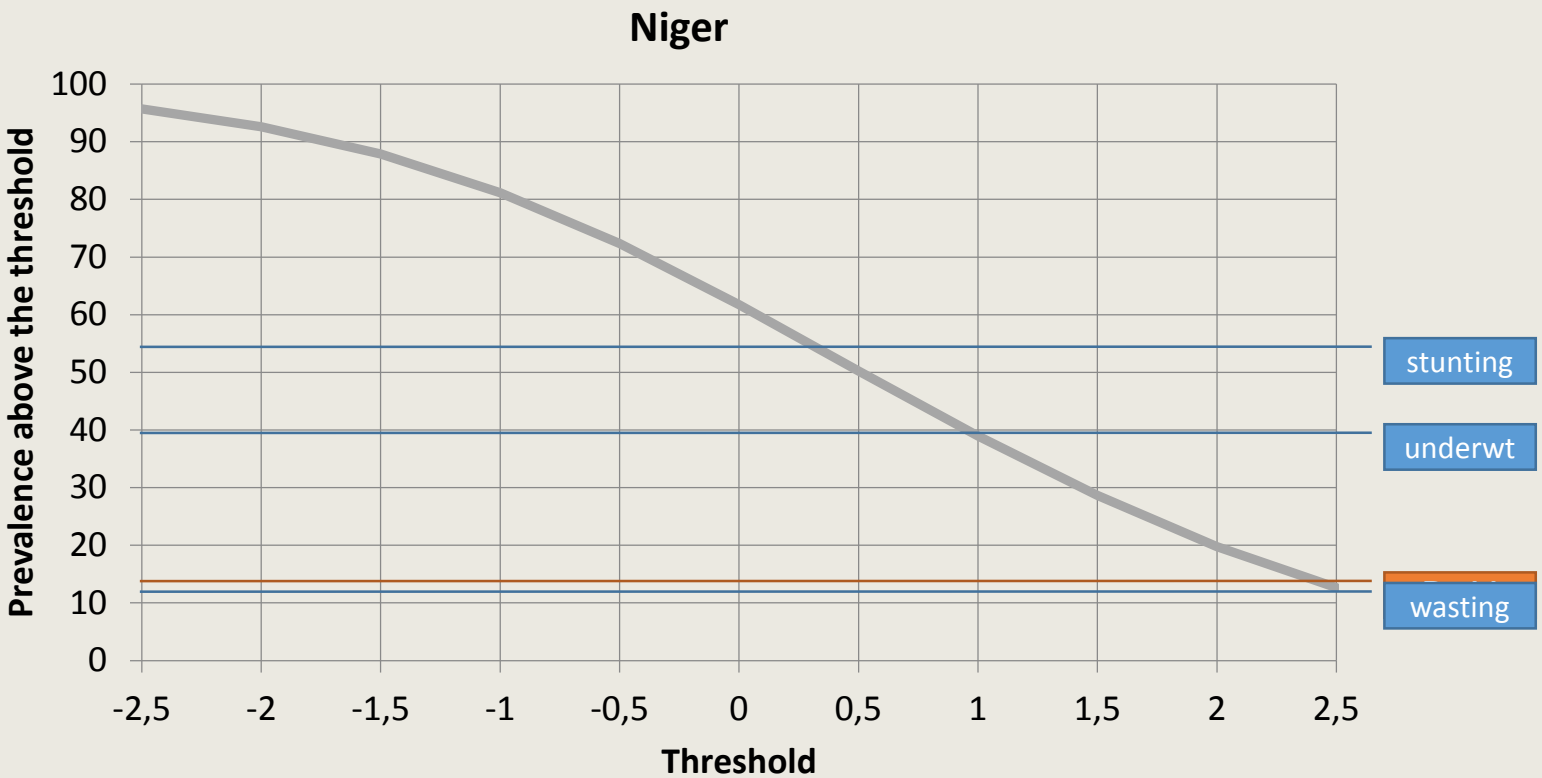
	PoU (2011-13)	Child anthropometrics (% among under 5)			
		Wasting	Underweight	Stunting	(2010)
Malawi	20.0	4.1	13.8	47.8	

Department for International Development

ROYAUME DE BELGIQUE Affaires étrangères, Commerce extérieur et Coopération au Développement

ESS

Prevalence rates at different thresholds



	PoU (2011-13)	Child anthropometrics (% among under 5)			
		Wasting	Underweight	Stunting	(2006)
Niger	13.9	12.4	39.9	54.8	

What's ahead?

- Global data collection in 2014
  - 150+ countries covered by the GWP
  - Face to face interviews in 112 countries
- Calibration of the global standards for common item severities and threshold settings
  - Data will be processed as they are made available from Gallup according to the GWP calendar
  - Once a sufficient number of countries from the various regions of the World are available, we will identify the severity of the common items, to be used as a reference for equating the scales and setting the thresholds
  - Scoring and corresponding prevalence rates will be determined in each country, considering both the common and the idiosyncratic items
- Two national level indicators based on prevalence rates
  - Food insecurity
  - Severe food insecurity (high risk of hunger)

Department for International Development
ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement
ESS

Thanks!

Terri Ballard
Carlo Cafiero
Pietro Gennari
Dana Glori
Anne Kepple
Mark Nord
Aymeric Songy
Sara Viviani
Verena Wilke

Voices-of-the-Hungry@FAO.org

Department for International Development

ROYAUME DE BELGIQUE
Affaires étrangères, Commerce extérieur et Coopération au Développement

ESS