2020 CONFERENCE PAPER 1 • MAY 2014

MEASURING RESILIENCE IN A VOLATILE WORLD

A PROPOSAL FOR A MULTICOUNTRY SYSTEM OF SENTINEL SITES

Christopher B. Barrett

Derek Headey

AUTHORS

Christopher B. Barrett (cbb2@cornell.edu) is the David J. Nolan Director and Stephen B. and Janice G. Ashley Professor in the Charles H. Dyson School of Applied Economics and Management, as well as Professor in the Department of Economics, Cornell University, Ithaca, New York, USA.

Derek Headey (d.headey@cgiar.org) is research fellow in the Poverty, Health, and Nutrition Division of the International Food Policy Research Institute (IFPRI), Washington, DC, USA.

IFPRI and its 2020 Vision Initiative appreciate the generous support of and active engagement with the consortium of partners for the 2020 conference.

In partnership with the African Union Commission, IFPRI and its 2020 Vision Initiative are proud to contribute to the 2014 Year of Agriculture and Food Security in Africa.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

IFPRI

WWW.2020RESILIENCE.IFPRI.INFO

2020 VISION

2033 K Street, NW, Washington, DC 20006-1002 USA | T. +1.202.862.5600 | F. +1.202.467.4439 | ifpri@cgiar.org | www.ifpri.org

.

CONTENTS

ABSTRACT

Substantial numbers of the world's chronically poor and malnourished population live in an increasingly volatile world. The dangerous nexus of climate change, rapid population growth, conflict, and food price volatility already appears to have pushed several poor regions into states of permanent crisis, even as the rest of the world has seen unprecedented progress against poverty. This disturbing state of affairs, along with our expanded knowledge of the intimate interactions between short-term shocks and long-run development, has catalyzed widespread interest in resilience building, and in what such a framework implies for understanding the causes and consequences of acute vulnerability to natural and man-made disasters.

In this paper we ask what this paradigm implies for the measurement and analysis of resilience. Resilience is fundamentally about complex dynamics. Slower-moving ecological, economic, demographic, and social stressors create vulnerability to short-run shocks, which in turn can have long-term consequences by reinforcing preexisting vulnerabilities. In our view, this basic conception of resilience has fundamental measurement implications. First, resilience can be measured and understood only through higher-frequency surveys that capture the causes and consequences of time-varying stressors and shocks, including seasonal shocks. Second, resilience can be understood only through surveys that capture the multidimensional complexity of stressors, shocks, and feedback loops, including the complex interactions between economic, social, and ecological forces. Third, the underlying stressors that create vulnerability, and the resilience-building interventions that reduce vulnerability, can be gauged and evaluated only over the longer term.

This conceptualization of resilience motivates us to go a step further than existing research on resilience and on food and nutrition security measurement, by outlining a far more expansive strategy for improving and scaling up the monitoring, measurement, and analysis of the world's most vulnerable populations. We propose the development of a multicountry system of high-frequency, long-term sentinel sites in the world's most vulnerable regions. If implemented along the lines we conceive, this system could be a high-return investment for resilience-building efforts, since it would serve multiple purposes. This system offers the only rigorous means of monitoring vulnerability and resilience in the world's most volatile regions. This system would bolster existing early-warning systems by complementing them with household-level indicators. This system would improve the targeting of emergency resources. This system would be instrumental for diagnosing the underlying sources of vulnerability, for identifying key thresholds of resilience, and for designing appropriate resilience-building strategies. And this system would provide a rigorous foundation for large-scale evaluations of resilience-building activities.

While there are strong justifications for such a system, the devil is necessarily in the details, and much of this paper is concerned with those details. Largely to learn from existing experience, we first review existing measurement strategies that are similar in purpose or design to the sentinel system outlined above. When implemented, long-term, high-frequency measurement systems have often yielded great benefits but been hampered by cost, lack of institutional coordination, and insufficient dissemination and usage of data. The need to keep costs down and benefits widespread therefore motivates us to consider which countries in the world have the highest priority for the development of sentinel sites, based on indicators such as child nutrition and health outcomes, exposure to disasters, and past emergency assistance levels from the international community. We then turn to crucial issues of data collection design by outlining a hybrid sampling and survey design that will help achieve the various objectives outlined above while keeping costs down. We also argue that the proliferation of mobile phones and other information and communications technologies offers substantial scope for a cost-effective system of this kind, far more so than would have been available in the past.

Finally, we consider who should lead and contribute to this ambitious effort. Since the principal advantage of this approach is that it can yield benefits for a wide range of institutions and purposes (relief and

development, operations and research, social and biophysical sciences), and since the costs of a long-term commitment to these sentinel surveys would be large indeed for any single agency, we propose the need for a relatively broad consortium of international donors. This consortium should first focus on establishing partnerships with national governments and then commit to long-term resilience monitoring as well as domestic capacity building. With this essential commitment in place, this consortium would then need to secure implementing partners with a permanent presence on the ground, as well as the technical expertise of international organizations of various sorts.

Ultimately, we argue, it is *only* this kind of long-term, cooperative commitment that will provide a scientific evidence base for diagnosing and resolving the world's worst problems of hunger, poverty, and malnutrition. *Only* this kind of sentinel system can generate the data and evidence needed to inform actions to build resilience and to help the global community eliminate extreme poverty in the generation ahead. The status quo is simply not enough.

Keywords: early warning, emergency response, humanitarian assistance, impact evaluation, resilience, risk, shocks, stressors, vulnerability, surveillance, poverty, malnutrition, food insecurity

ACKNOWLEDGMENTS

This paper would not have been possible without the extensive and rich insights of a large number of people who have worked on measurement challenges for a number of years. From Helen Keller International, we thank Nancy Haselow, Victoria Quinn, and Erica Khetran. We thank researchers currently or formerly associated with Cornell University, especially Suresh Babu, Mark Constas, Jean-Pierre Habicht, David Pelletier, Per Pinstrup-Andersen, Wendy Wolford and Dan Sisler. From the World Food Programme, we thank Kathryn Ogden, Astrid Mathiassen, and Jean-Martin Bauer. A number of colleagues at the International Food Policy Research Institute (IFPRI) and Cornell University also provided useful comments and ideas, including Rajul Pandya-Lorch, Marie Ruel, John Hoddinott, Stuart Gillespie, and Nathaniel Jensen. Finally, a range of other colleagues provided comments on the shorter policy brief version of the paper, including Dan Maxwell, Greg Collins, Tim Frankenberger, Tiffany Griffin, Tim Mahoney, Gero Carletto and Simon Levine. We strongly emphasize that their very helpful conversations in no way implicate these colleagues in our arguments, assumptions, and recommendations, responsibility for which is solely our own. We also extend our thanks to Carlo Azzarri and Melanie Bacou for providing us with maps of Ethiopian child nutrition outcomes and agroecological zones.

1. INTRODUCTION

At no time in modern history have high-level public- and private-sector leaders paid more attention to the related but distinct objectives of reducing poverty, hunger, and malnutrition, and ameliorating the impacts of natural and man-made disasters that too regularly strike low-income communities around the world. As such interest has grown, so has our understanding of these problems, especially their complexity and interconnectedness. We know that poverty encompasses the chronically poor as well as those who suffer transitory poverty due to economic, sociopolitical, or weather shocks, or who become trapped in long-term destitution following severe disruptions to their livelihoods (Carter and Barrett 2006; Barrett and Carter 2013). We know that malnutrition in early childhood, and even in utero, can have lifelong consequences for physical, cognitive, and economic development (Hoddinott et al. 2008; Victora et al. 2008; Almond and Currie 2011). We know that hunger and malnutrition have many causes that go well beyond inadequate production of food, to lack of access and entitlements (Sen 1981; Devereux 2006). And we know with increasing certainty that the poor live in an ever more volatile world and will likely face more frequent and more severe climatic shocks that will have long-term consequences on their ability to permanently and rapidly escape poverty, hunger, and malnutrition (IPCC 2012).

This knowledge, though still far from complete, has recently prompted the international community to embrace a pro-poor concept of resilience building, particularly in the context of food and nutrition security (Béné et al. 2013). The emergent resilience paradigm integrates long-standing interests in early warning from humanitarian assistance; in empowerment and poverty reduction from development assistance; and in risk management and systems linkages from food-security studies, ecology, and climate change science. Resilience offers not so much a new set of ingredients as a new recipe for combining them so as to more effectively understand the complex and interconnected relationships between the stressors people face, the shocks they suffer, and the time path of recovery (or nonrecovery) in key outcome indicators related to human and system well-being. Toward that end, resilience is increasingly understood as a capacity that prevents individuals, households, and communities from falling below a normatively defined standard of living, whether defined in terms of poverty, health or nutritional status, subjective well-being, or some other measure (Barrett and Constas 2013; FSIN 2013).

The resilience paradigm therefore offers three key conceptual strengths to policymakers, development and relief agencies, and researchers. First, resilience incorporates fundamentally important *dynamic* concepts—the dynamics of moving in and out of poverty, hunger, and malnutrition—that few other paradigms can. Second, it emphasizes the central role that *risk* plays in human lives, both the stressors people face on a daily basis (the risk of something bad happening) and their experience of shocks (bad things actually happening). Third, the resilience paradigm focuses our attention on interconnected social, economic, and ecological *systems*, and on the way that these subsystems critically interact with each other.

Yet the richness of resilience as an organizing framework poses tremendous challenges for our ability to measure and monitor resilience and to evaluate the impact of interventions intended to build resilience. This paper focuses on how the global community might overcome these challenges. This is far from an arcane academic issue. Good measurement drives accurate diagnosis as well as timely and appropriate response (Barrett 2010), and it fuels the learning required to rediagnose, to evaluate and reevaluate, and to continuously improve response. To date, however, widespread enthusiasm for the concept of resilience has not been matched by significant progress in measuring resilience, although the Food Security Information Network (FSIN) is now mounting a concerted, coordinated effort on that front (FSIN 2013). Moreover, the development community has made great strides in measuring and understanding secular changes in poverty, malnutrition, and hunger in recent decades, and there are clearly important lessons to be learned from these efforts (see Section 2).

The considerable progress achieved over the past generation in the measurement of poverty, health status, and ecological conditions, along with considerable progress in the technical underpinnings of early-warning systems and emergency food-security assessments (see Section 3), provides a wealth of knowledge on which the global community can build a systematic approach to measuring and monitoring resilience indicators. But the existing systems do not themselves satisfy central demands that arise from the three key conceptual strengths of resilience: dynamics, risk, and systems. The array of increasingly high-quality household surveys in developing countries leaves us with a major data and knowledge gap in the context of resilience-building efforts.

Defined as it is by a much sharper focus on altering the dynamics of human welfare, resilience building requires a very different empirical strategy focused on ***high-frequency***, ***sustained***, ***long-term*** surveying of a network of carefully selected ***sentinel sites*** (hereafter "sentinel system," for short).[1] Resilience data must be collected at high frequency in order to capture the impacts of stressors and shocks (and responses to shocks) using risk-sensitive indicators. The data must be collected over the long term, because vulnerability to shocks is the product of slower-moving stressors as well as of long-term, multisectoral interventions for building resilience. The data should be collected in sentinel sites that are strategically selected for the purposes of minimizing costs while maintaining representativeness of key structural characteristics, such as specific agroecologies or livelihood zones. We also make a concrete proposal that a multicountry system of sentinel surveys should be implemented by a consortium of international agencies, nongovernmental organizations (NGOs), and national governments.

The costs and challenges associated with developing and maintaining such a system will be substantial. Much of this paper is therefore concerned with documenting these obstacles and proposing means for overcoming them. Yet we also maintain the view that the benefits of such a system will be immense and multidimensional. These benefits are as follows.

First, *only* this kind of system can accurately monitor long-term changes in resilience, defined generally along the lines of "a capacity that prevents individuals, households, and communities from falling below a normatively defined level for a given developmental outcome (e.g., food security, poverty level, well-being)" (FSIN 2013, 7). Infrequent household surveys used for monitoring poverty or health status at national scale are a poor tool for gauging the impact of occasional but large-scale covariate shocks, such as droughts or floods, or of more individual- or household-specific shocks, such as illness or unemployment, as well as of year-to-year fluctuations associated with seasonality. Climate change, environmental degradation, and population growth all appear to increasingly expose poor populations to natural hazards while the spread of small arms and political instability expose them to man-made risks. Even as the world enjoys considerable progress in the struggle against poverty and hunger globally, the increasingly complex and vexing challenges in those regions left behind necessitate a well-designed and well-implemented sentinel system to monitor long-term changes in resilience and to evaluate the effectiveness of different interventions.

Second, while early-warning systems based on climate, market, and administrative-level data have improved significantly over the past generation, a high-frequency survey of *people* offers substantial scope to improve the efficacy of early-warning systems. Individuals and households themselves (and by aggregation, communities) can offer a bottom-up perspective of imminent threats by reporting on their grain stocks, on the health and nutrition of their families and their livestock, on the various coping strategies that they invoke in the face of shocks, and on their subjective impressions of what the near future holds. Vast technological improvements in mobile communication, data collection, and data processing mean that these household-level indicators could be aggregated, cleaned, and disseminated in near-real time.

[1] It is perhaps worth reminding readers of the definition of *sentinel*: one that stands guard; from the Latin, *sentina*, for watchfulness.

Third, the construction of a large system of sentinel sites within a relatively small number of developing countries that are both exposed and vulnerable to major covariate shocks might help to mobilize and target resources more effectively, both across and within countries. Our perception is that the current means of channeling emergency resources to vulnerable populations is considerably improved over that of a generation ago but still far less evidence based and systematic than it could and should be. In some contexts, for example, governments and international development partners have starkly disagreed about the magnitude of a crisis. But just as important is the immense challenge of defining and targeting needy subpopulations, especially in regions where a majority of the population is chronically poor and vulnerable, and international resources are limited. The current approach to assessing emergency needs offers surprisingly little in the way of rigorous, reliable, and objective data to zero in on the truth of a context.

Finally, *only* this kind of multidimensional, high-frequency, long-term sentinel system can enable rigorous program impact evaluation of resilience-building efforts conducted at scale, as well as a more thorough understanding of the complex nexus between the economic, social, and ecological forces that determine resilience. The current vogue of small-scale project evaluations through randomized controlled trials (RCTs) typically only achieves (at best) internal validity, but can potentially suffer from a serious lack of external validity and sustainability. Evaluating programs at scale therefore requires much larger survey instruments, typically panel surveys (Elbers and Gunning 2013), as well as long-term surveys to establish sustainability. Longer-term surveys will also be essential for understanding the interactions between economic, social, and ecological stressors. RCTs and other experiments can be readily nested within the sentinel system to facilitate impact evaluations where experimental methods are feasible and appropriate.

We believe these are compelling arguments for investing in a new system of sentinel sites for resilience measurement. But acknowledging that the devil is in the details, the remainder of this paper examines those details. In Section 2 we delve more into *why* we need such a system. We first discuss what resilience means, why this paradigm adds value to existing constructs, and what the resilience paradigm implies for measurement. In Section 3 we briefly review existing measurement strategies, including those that bear close affinities to the system we have in mind. In Section 4 we look at *where* such a system might be most appropriately located by first focusing on country-level vulnerabilities, before outlining and illustrating how sentinel sites need to be strategically placed across different spheres of vulnerability, livelihoods, and ecologies within chosen countries. In Section 5 we briefly discuss *what* to measure and *how* to measure it.[2] We discuss the delicate balances that need to be attained in order ensure that this system can achieve multiple objectives. In Section 6 we focus on potential institutional arrangements and implementation challenges. Experience tells us that there are four interrelated challenges: long-term commitment, interagency coordination, capacity, and ensuring sufficient usage of data.

Any one of these constraints could limit the success of a new and more ambitious measurement system for food and nutrition security. But while we should not be naïve in our expectations of what can be achieved by such a system, we must also bear in mind that the status quo approach is nowhere near good enough. No existing system was developed with the objective of monitoring resilience and none lends itself readily to suitable modification, although existing systems offer a great many valuable lessons that must inform the design of a sentinel system for resilience building. Many of the regions in which sentinel systems are most necessary already experience "permanent crises"—a term often used with no hint of irony in the humanitarian relief community—as well as sharp seasonal fluctuations in food availability and individuals' nutritional status. Especially if the dangerous nexus of climate change, rapid population growth, and sociopolitical instability leads to even greater exposure to both natural and unnatural disasters in the years to come, as is widely

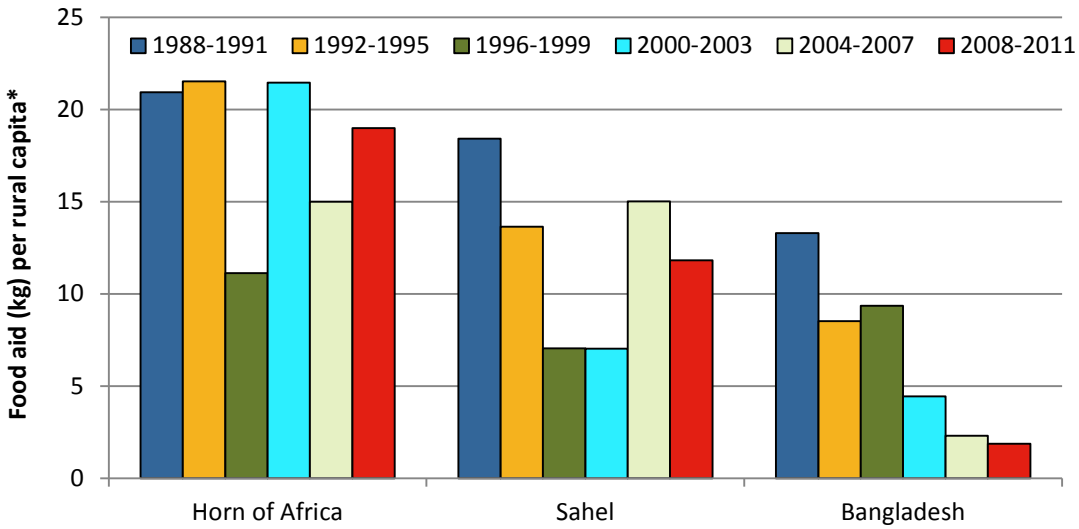
[2] The FSIN technical series from the Resilience Measurement Technical Working Group promises more detail in this important area.

expected, then building resilience effectively will require major improvements in our ability to monitor and understand the world's most vulnerable populations. More precisely, we are convinced that the world needs a sentinel system designed for the purpose of measuring and monitoring resilience and evaluating the impact of interventions intended to build resilience.

2. THE RESILIENCE PARADIGM AND ITS IMPLICATIONS FOR MEASUREMENT, MONITORING, AND ANALYSIS

Much of the recent surge of interest in resilience stems from concern that the most vulnerable regions of the developing world face recurring and ever more severe crises, and that these crises stem from the complex interaction of economic, social, political, demographic, and ecological factors (Barrett and Constas 2013). Not only has the number of people affected by natural disasters and conflict been increasing steadily, but the most severely affected are also increasingly concentrated in a modest number of fragile, failed, or failing states. Figure 2.1, for example, shows trends in food aid receipts—a reasonable proxy for lack of resilience at the national level—in the Horn of Africa, the Sahel, and Bangladesh over the period 1988–2011. While Bangladesh has seemingly had tremendous success in resilience building in recent decades, trends in exposure and vulnerability to disasters for the Sahel and the Horn of Africa remain constant over time and deeply worrying given the threat of further climate change (Headey, Taffesse, and You 2014).

Figure 2.1 Trends in food aid receipts in the Horn of Africa, the Sahel, and Bangladesh, 1998–2011



Source: Authors' estimates from WFP (2013).
Note: * We assume that the vast majority of food aid goes to rural people even though the WFP (2013) data pertains to national level populations. Nevertheless, these regions have roughly similar rural population shares, and using total population in the denominator yields very similar trends.

In such places especially, reconciliation of humanitarian relief in response to emergencies with longer-term development efforts has long proved difficult. Concerted efforts to build resilience are widely seen as a best-bet approach to realizing latent synergies and to achieving sustainable balance between emergency and development assistance. In this section we first revisit the definition and value-added of resilience as a concept. We then explore what this conceptualization implies for measurement. Finally, we illustrate how most existing measurement systems fall well short of meeting the essential criteria for resilience, although there are important lessons to be learned from existing approaches to nutrition surveillance, emergency food-security assessments, early-warning systems, and long-term ecological resilience programs.

WHAT IS RESILIENCE?

Resilience in the context of development is an umbrella under which a selection of existing research, policy, and operation activities may be usefully organized. The challenge of any umbrella term is to define what it is that falls under the umbrella, and what it is that does not. While no clear consensus yet exists in favor of any of the many proffered definitions of resilience in the recent literature (see Béné et al. [2012] for a review), we employ a recent definition proposed by Barrett and Constas (2013), which has been adapted by the multiagency Resilience Measurement Technical Working Group (FSIN 2013):

> Development resilience is the capacity *over time* of a person, household or other aggregate unit to *avoid poverty* in the face of various *stressors* and in the wake of myriad *shocks*. If and only if that capacity is and remains high, then the unit is resilient. (Barrett and Constas 2013, 3, emphasis added)

This definition is attractive for four reasons.

First, it implicitly emphasizes that resilience is a stochastic and dynamic concept: building resilience means reducing the likelihood and severity of negative life events *over time*.[3] This definition therefore encompasses both chronic ill-being and transient ill-being caused by shocks. Barrett and Constas (2013) further emphasize that a central feature of the resilience paradigm is its ability to incorporate (but not require) complex nonlinearities and multiple equilibria in well-being measures—commonly associated with poverty traps—as well as vicious and virtual feedback cycles and ecological tipping points, where such features exist.

Second, this definition of resilience explicitly recognizes the role of background risk ("stressors") of various sorts and that sometimes risk turns into adverse events ("shocks") that can catastrophically change lives. Many definitions in the literature focus solely on shocks as the superficial, proximate cause of food and nutrition insecurity. But the aggregate adverse impact of exogenous hazards such as climatic shocks or violence turns as much on the preemptive responses in which people engage when faced with prospective shocks as it does on the adverse realization of random events.[4] Furthermore, multiple stressors and shocks occur simultaneously and interact in complex ways (for example, climate change and population growth are synergistic stresses). Resilience relates to capacity to maintain well-being in the face of any of a range of anticipated or unanticipated shocks or stressors, be they biophysical, economic, or sociopolitical in origin. Too many discussions of resilience focus only on the threats posed by a single subset of the myriad stressors and shocks that vulnerable communities, households, and individuals confront. This broader definition can encompass resilience to any specific shock or stressor while simultaneously accommodating aggregate, omnifactorial exposure.

Third, this definition distinguishes between different levels of aggregation: individuals, households, and communities (as, admittedly, do others in the literature). One potentially distinguishing feature between the resilience paradigm and at least some related antecedents is that the burgeoning resilience literature has focused substantial attention on the interactions between different spheres of an economic system, particularly between households and communities, and between communities and their ecological environments. Building on an extensive ecological literature on resilience that focuses on systems and the feedback among subsystems, resilience emphasizes multilevel interactions and multiple scales of analysis (Barrett and Swallow 2006; FSIN 2013). In contrast, much of the traditional economic literature on poverty or food security has focused solely on household- or individual-level analyses.

Fourth, unlike parallel literatures on resilience in ecology, engineering, or geography, this definition focuses squarely on human well-being outcomes measured against some normative standard—such as poverty or food insecurity. When resilience is defined as the capacity to avoid poverty (or food insecurity or some other indicator of well-being), mere return to an unacceptable status quo ex ante is insufficient. Thus resilience should not be equivalent to system stability when the systems in question are characterized by widespread human suffering.

The resilience paradigm therefore encompasses a focus on complex human welfare dynamics that is especially appropriate for contexts in which chronically poor and vulnerable subpopulations confront varied and prominent economic, social, and ecological stressors as well as frequent and intense exposure to shocks.

MEASURING RESILIENCE

The preceding brief reflections on the scope of resilience also define the scope of the measurement challenge. As our introduction noted, good measurement is crucial for diagnosis of problems; for mobilizing and targeting short-term resources; and for designing, implementing, and evaluating appropriate long-term resilience-building strategies. But what does good measurement mean in the context of concepts such as risk;

---

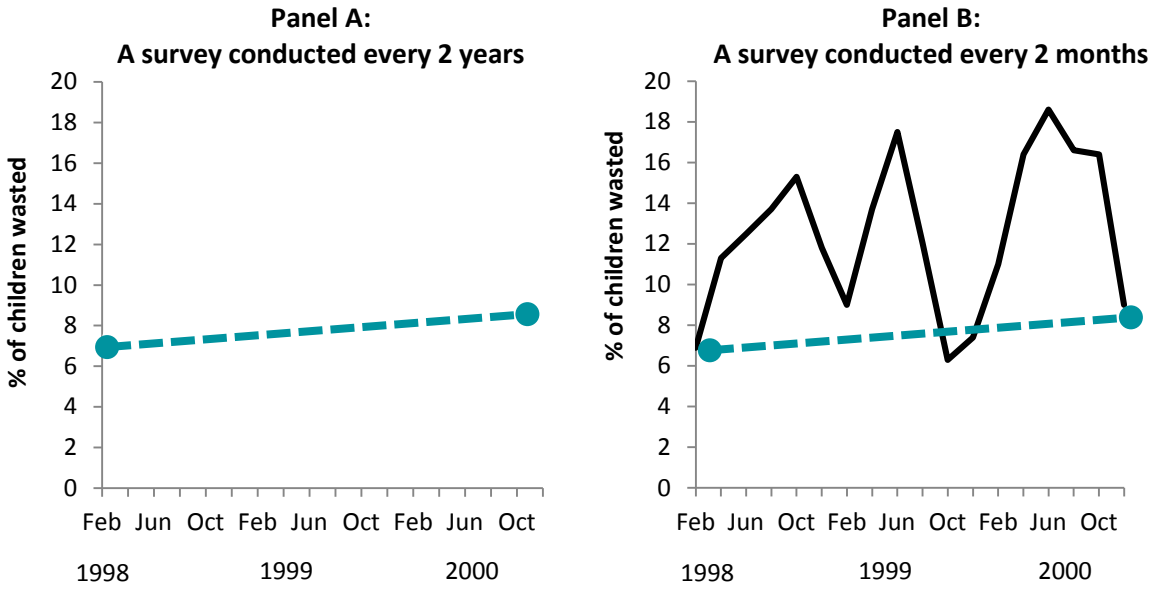
[3] The most common internationally agreed-upon definition of food security similarly emphasizes that food security exists when "all people, at *all times*" have "access to sufficient, safe and nutritious food" (FAO 1996, n.p., emphasis added).

[4] See Rockmore (2013) for an example of the economic impacts of violence in northern Uganda.

vulnerability; chronic and transient poverty; poverty traps and thresholds; and complex interactions between shocks and stressors, and between households and their social, economic, and ecological environments?

First, and most importantly, resilience measurement *must* involve ***sustained, higher-frequency surveys*** than are currently standard in most of the social sciences (Headey and Ecker 2013; Barrett and Constas 2013). This is because resilience relates to the likelihood over time of avoiding poverty, hunger, or malnutrition. While short intervals between surveys can overstate mobility by giving excessive weight to measurement error and transitory phenomena (Naschold and Barrett 2011), longer intervals between surveys can also obviously give very deceptive impressions of resilience. Figure 2.2 strikingly demonstrates this point with high-frequency data from the Nutrition Surveillance Program (NSP) in Bangladesh (Bloem, Moench-Pfanner, and Panagides 2003). Panel A reports observations of child wasting from two surveys, one conducted in February 1998 and the other in October 2000. The figure shows that child wasting was essentially unchanged over this period and relatively low: 8 percent in the 1998 round and 9 percent in the 2000 round. The second panel instead shows bimonthly data from the same NSP survey instrument. These high-frequency data tell a very different story: from troughs of 6–8 percent in the December to February period, child wasting doubles every monsoon (June to August) to levels of 15–18 percent. The magnitude of this severe seasonal variation is clearly not discernible via infrequent snapshots, or the more qualitative approaches to vulnerability assessments favored by many international agencies.

Figure 2.2 The difference between low- and high-frequency measurement: The example of child wasting in Helen Keller International's Bangladesh Nutrition Surveillance Program



Source: Authors' construction from data presented in Bloem, Moench-Pfanner, and Panagides (2003).

Second, resilience needs to be ***measured repeatedly*** over the ***longer term***. As the definition above notes, resilience is not just about shocks and their impacts but also pertains to understanding the longer-term stressors that create vulnerabilities to shocks. These underlying stressors cannot be reliably measured and analyzed in a three-year survey. Resilience is also defined as the ability to withstand or bounce back from shocks. But shocks are unpredictable. A region historically vulnerable to shocks may suddenly have several years—even a decade—of favorable weather, such that household and community resilience temporarily remains untested.[5] Furthermore, measurement error inevitably contaminates any single set of observations but can be satisfactorily addressed with sufficient repeated observations of the same respondents or

[5] Arguably, southern African countries such as Malawi are now experiencing a long run of relatively favorable weather, after a severe drought in the early part of this century.

populations (Griliches and Hausman 1986). Finally, resilience-building interventions need evaluation at scale over the longer term, long after the interventions have concluded, in order to establish reliably whether they have intended (or unintended) effects and whether such effects persist postintervention. The past decade has witnessed a burgeoning literature on experimental project evaluation in economics and elsewhere in development studies; but too little of this literature has focused on long-term sustainability. There are many reasons why projects might have favorable short-term impacts but declining impacts over the longer term, and vice versa (for example, if social learning is important, or if monitoring effort weakens over time).

Third, resilience indicators should include ***sufficiently sensitive indicators*** of stressors, shocks, coping strategies, and human welfare. In the nutrition sciences, for example, it is well documented that weight for height, mid-upper-arm circumference, and adult body mass indexes are appropriately sensitive indicators of acute malnutrition. But while the last two decades have seen a tremendous amount of experimentation with food-security indicators, surprisingly little of that research has focused on the sensitivity of indicators to shocks. One exception is Headey, Ecker, and Trinh Tan (2014), who compared and contrasted calorie indicators, dietary diversity indicators, and subjective indicators. They found that economic theory and empirical experience strongly suggest that dietary diversity indicators—such as the World Food Programme's (WFP's) food consumption score—ought to be sufficiently sensitive to sizable shocks, but they warn against naïve overreliance on largely untested subjective indicators. Subjective measures can nonetheless be extremely valuable for eliciting respondents' perceptions of stressors and other uncertain future events (Manski 2004; Doss, McPeak, and Barrett 2008; Delavande, Giné, and McKenzie 2011). Other likely shock-sensitive indicators of relevance would include coping strategies, prices and wages, food stocks, and animal health and nutrition indicators, several of which are rarely included in the workhorse survey instruments on which most poverty and food-security measurement relies.

Fourth, resilience needs to be ***measured at multiple levels:*** for individuals (particularly nutritional status), households, and communities, but also for local economies and ecologies. Currently there is an unnecessary disconnect in the analysis of these different units. Ecological research has only relatively recently started to integrate the study of human well-being with that of the dynamics of nonhuman organisms and subsystems.[6] Economic and social research has likewise been relatively weak at incorporating ecological and climatic factors into analysis. A broader measurement strategy that combines individual- and household-level data with community-level data (including qualitative data), macroeconomic data, and a variety of ecological variables would help create the space for more innovative and robust interdisciplinary analysis of vulnerability and resilience.

Fifth, while there are many advantages of common measurement strategies across different contexts (particularly in terms of using surveys for targeting resources and for comparative analysis and program evaluation), there will also be some need for ***context-specific measurement***. Different regions of the world vary across all the domains of interest: initial vulnerabilities, livelihoods, and asset portfolios; the types and magnitudes of shocks and stressors; the policy and institutional arrangements; and the broader social, economic, and ecological environments they face. A well-established approach is to develop a core set of template data collection instruments and methods so as to ensure reasonable comparability across sites, and then adapt each instrument and method to the context so as to tailor it to the population of interest.[7]

The measurement challenge raised by the emergence of resilience as an organizing framework for a wide range of agencies demands that detailed technical attention be paid to the specific methods to be deployed following the principles outlined above. This is not the place to explore those details, but the FSIN's Resilience Measurement Technical Working Group is currently tackling this question.[8]

---

[6] See, for example, the August 23, 2011, special feature of the *Proceedings of the National Academy of Sciences* titled "Biodiversity Conservation and Poverty Traps."

[7] Examples include the Demographic and Health Surveys financed by the United States Agency for International Development and the Living Standards Measurement Surveys run by the World Bank.

[8] For details see www.fsincop.net.

3. RESILIENCE MEASUREMENT IN PRACTICE: EXISTING APPROACHES AND A NEW PROPOSAL

Having broadly outlined how we think resilience should be defined and measured, we now consider how existing measurement approaches go about it, before outlining a new and admittedly ambitious proposal for a more far-reaching and systematic approach. There is no reason to reinvent the wheel, and many aspects of current measurement and monitoring systems work very well and merit replication. At the same time, one should be especially careful not to reinvent the flat tire, so we want to take care to learn lessons from less successful experiences, or even from the weaker elements of successful data collection systems.

HOW WELL DO EXISTING APPROACHES MEASURE RESILIENCE?

To what extent do existing measurement approaches fulfill the measurement principles outlined in the previous section? A variety of organizations regularly field a broad range of surveys in the developing world. These can broadly be categorized into economic surveys, health and nutrition surveys, and what we would term food and nutrition security surveys. Standard economic surveys, such as the World Bank's Living Standards Measurement Surveys (LSMS), the workhorse source of global poverty statistics, have yielded many rich insights in recent decades, and they certainly possess some of the requisite characteristics described above. But the infrequent measurement of these surveys (typically every 4–5 years) is a deal breaker as far as resilience measurement is concerned. The same is true of the excellent Demographic and Health Surveys (DHS), which have greatly expanded the availability and analysis of child and maternal health and nutrition data. All of these surveys, and those like them, are simply not geared toward measuring the impact of shocks, stressors, and feedback across subsystems, especially at high frequency. Thus, they clearly fail to meet the near-term programming needs of humanitarian agencies, nor do they meet the longer evaluation requirements of resilience-building efforts.

The surveys most closely aligned with the objective of gauging shocks and stressors are those we term "food and nutrition security" surveys. These cover instruments like the WFP's Food Security Monitoring Surveys (FSMS) and less common, higher-frequency "surveillance" or "sentinel" surveys conducted by Helen Keller International (HKI), UNICEF, Cornell University, and various NGOs and developing-country governments.[9] Two surveys that come close to meeting the measurement principles outlined above—and that illustrate both the usefulness of this strategy and the key issues that demand attention in order to avoid expensive failure—are the Arid Lands Resource Management Programme (ALRMP), which has collected data in the most drought-prone regions of Kenya from 1988 to present, and HKI's Nutrition Surveillance Program in Bangladesh, which collected data from 1990 to 2003 and again from 2009 to present. Both of these surveys are long term, high frequency, and clearly focused on measuring and understanding the causes and consequences of household- and individual-level shocks and stressors. More recently, the WFP has also implemented FSMS in a number of countries. While typically only biannual surveys, and often more narrowly focused on monitoring food security than assessing underlying causes, the FSMS constitutes a significant improvement over the use of one-time surveys and vulnerability assessments in vulnerable environments.

What are the lessons from these approaches?

The ALRMP has achieved some success in quite difficult environments and in spite of a raft of problems with quality control in data collection and entry (Johnson and Wambile 2011). ALRMP achieved longevity,

[9] See McKinney (2011) and Shoham, Watson, and Dolan (2001) for reviews. It is important to note that "nutrition surveillance" refers to quite a broad range of measurement strategies, including basic nutritional monitoring in which little or no socioeconomic background information is collected. Because these systems typically have narrow programmatic objectives—such as identifying and treating children with severe acute malnutrition—we do not consider them here. Our focus is instead on multipurpose surveys.

sustainability, and increased spatial coverage over time, from just 1 district to 28 districts, in more arid pastoralist and agropastoralist regions of Kenya. The system is fully integrated into national early-warning and drought response decisionmaking processes. The resulting strong demand for the data by decisionmakers has been central to sustaining the ALRMP data-collection apparatus. ALRMP also continues to supply data that have been used to identify predictor indicators for deteriorating nutritional status (Mude et al. 2009) and to develop insurance products to help protect vulnerable populations against drought (Chantarat et al. 2013). But the usefulness and reliability of the data have been limited by underfunded data collection and entry supervision, which have especially limited the use of ALRMP data for impact evaluation and broader research (Johnson and Wambile 2011).

HKI's Nutrition Surveillance Program (NSP) in Bangladesh was perhaps even more successful in terms of generating high-quality data for monitoring and analyzing vulnerability. With support from the United States Agency for International Development (USAID) and later the Dutch government, HKI initiated the NSP in 1990 after the devastating floods of 1988 highlighted the lack of reliable population-based data to target assistance to the affected areas. It was determined that there was a need to be able to assess the impact of the many natural disasters in Bangladesh on the nutritional status of children and women through an established system to measure current, predisaster, and postdisaster situations, in comparison with nonaffected areas. Indeed, from 1990 to 1997 the NSP was not nationally representative but instead specially targeted toward disaster-prone areas. The NSP was then made nationally representative and continued in that form until 2003, when the program ended. In 2009 a more streamlined NSP was reinitiated, this time in partnership with BRAC (formerly the Bangladesh Rural Advancement Committee). The 1990–2003 NSP was designed as a panel of villages, generating repeated observations over time on the village as a whole, albeit not on individuals within the survey villages. The sampling strategy changed somewhat with the 2009 NSP reintroduction, to a panel of multivillage zones.

The NSP was innovative in a number of regards (Akhter and Haselow 2010; Bloem, Moench-Pfanner, and Panagides 2003; Shoham, Watson, and Dolan 2001). First, it was unusually well funded, though arguably at a level that made sustainability and replication infeasible (see Section 5). Second, the data collection process was quite unique. Field staff from 26 national NGOs constituted a permanent presence in the sentinel sites, and a wide range of data were collected (on child anthropometry, household expenditures, coping mechanisms, and other health and demographic variables) on a bimonthly basis in order to coincide with Bangladesh's quite idiosyncratic climate, which was believed to heavily influence nutrition outcomes (Figure 2.2). These data were also subject to strong quality control measures.[10] Third, there was a strong focus on data usage, including both dissemination of NSP bulletins and some strikingly innovative research (see, for example, Torlesse, Kiess, and Bloem [2003] and, in the Indonesian context, Block et al. [2004] on the impacts of macroeconomic shocks on nutrition). Last and not least, the NSP appears to have been extremely useful in driving diagnosis and response. In August 1998, one of the worst floods of the century occurred. While the NSP had been collecting flood data for years by then, the flexibility of the system allowed the addition of several questions to the questionnaire to capture information on the extent and magnitude of this particularly devastating flood. The permanent presence of well-trained survey enumerators and data managers meant that an existing system was in place to collect, process, and release the data remarkably quickly, which allowed local NGOs to advocate for and secure resources for flood-affected households. Thus, individual- and household-level data were able to extensively supplement climate-based early-warning systems and to say something about the human consequences of the crisis, rather than just the extent of the biophysical shock.

[10] HKI/Bangladesh implemented the NSP in collaboration with the Institute of Public Health Nutrition and a total of 26 local NGOs. Shoham, Watson, and Dolan wrote that "in this manner, HKI/B has not only established a viable information system but also a system for transferring its expertise in conducting nutrition surveillance to local counterparts in support of building local capacity for the future" (2001, 34).

Nevertheless, HKI's surveillance systems in Bangladesh and Indonesia reportedly still faced significant challenges, including insufficient government participation and use of data. HKI's new system in Bangladesh has addressed these challenges but now faces the additional challenge of struggling to respond to high levels of demand with appropriate and timely dissemination products.[11]

The WFP's FSMS are a relatively recent innovation. Traditionally, the WFP had a relatively ad hoc approach to food-security assessment based on contextual factors. In many emergencies the WFP uses what it terms "initial assessments" and "rapid assessments," in which it relies on key informant interviews, focus group discussions, secondary data, and strong assumptions to assess and target vulnerable populations. In countries in which it has a more permanent presence it also conducts household surveys, though these are often one-time surveys used to broadly map vulnerability and to understand its basic correlates. Hence, the more recent adoption of the higher-frequency FSMS offers substantially more scope to quantify and analyze vulnerability in shock-prone regions, but lack of funding for higher-frequency surveys reportedly remains a serious constraint.

These examples show what can be achieved with long-term commitment, sufficient planning and training, and clear strategies for data dissemination and uptake to ensure that end-user demand is built in from the outset. But these examples also hint at important limitations that would need to be addressed with the introduction of a network of sentinel sites. Moreover, it is worth asking why other higher-frequency surveys have not been self-sustaining, or have been resisted from the outset. Why do relief agencies such as the WFP and its NGO partners not engage in high-frequency collection as a norm, rather than as an exception?

The first problem we perceive is high costs, though this is very much related to "market failures" or coordination problems and the fact that data of this nature are in large measure a public good. Long-term, higher-frequency surveys are obviously expensive. HKI's very innovative systems in Bangladesh and Indonesia were unusually expensive. The more streamlined NSP currently being implemented in Bangladesh by HKI and BRAC is far more sustainable, at around US$1 million[12] per year. But an analogously sized survey in Africa would face much higher transportation costs, and quite likely higher labor and training costs (even though most African countries are much less populous than Bangladesh). These costs are obviously sizable for a national government to commit to in addition to the many other economic, health, and nutrition surveys it conducts. So donor funding and technical oversight have been the norm in most of these surveys.

But underlying this challenge are the lack of coordination and the related nature of data as a public good:[13] a well-designed, long-term, high-frequency survey would benefit public and private agencies working in both relief and longer-term development. While the costs of coordination typically make multiagency implementation of data collection and analysis cumbersome, if any single agency were asked to bear the full financial and logistical burden of mounting such an exercise, then the costs of higher-frequency and multidimensional surveys would typically be excessive and cause suboptimal supply, if not preclude ever undertaking the effort. We therefore perceive that lack of coordination is the central underlying problem rather than sheer monetary costs. Given global overseas development and emergency assistance budgets in excess of $100 billion annually, a price tag of $1–2 million per year per country included in the sentinel system is clearly manageable, especially if it appreciably improves the cost-effectiveness of other expenditures. The core issue is therefore coordination among donors, governments, and implementing agencies.

A second widely cited problem, with nutrition surveillance systems in particular, is lack of country ownership and a related lack of data usage demand. Reviews of health and nutrition surveillance systems from

---

[11] We thank Erica Khetran of HKI for this insight.

[12] All dollar amounts are in US dollars.

[13] Technically, data are a "club good," meaning they can be used by multiple parties without degradation (the "nonrival" criterion of a public good) but access to data can be controlled (so they do not satisfy the "nonexcludable" criterion). But club goods still face the coordination problem referred to here, so we use the more familiar "public good" shorthand for present purposes.

earlier decades often concluded that too little thought was given to strategies for data and research dissemination, especially for policymakers (Babu and Quinn 1994; Jonsson 1995; McKinney 2011). The donor-driven nature of some of these surveys exacerbated these problems.[14] Effective data collection systems need to be designed from the outset with an eye to who will use the data and how. The sorts of questions around which systems need to be designed are these: Whose actions might be influenced by the data—or by analyses derived from the raw data, the processed data, or both—and how? When and in what form do data or analyses need to reach decisionmakers in order for the information to have value? A demand-driven approach requires designing systems starting from intended decisions that might be better informed by data, and then working backward to the necessary analyses and information products, the necessary data processing and collection, and the survey instrument design and sampling strategies. Put differently, the policy options that can be influenced by evidence must inform analysis, which must in turn drive measurement if one aspires for measurement and monitoring systems to ultimately guide policy, thus completing the loop.

Finally, the nature of poor and highly vulnerable environments means that logistical and human technical capacity will often be a constraint, at least in the short term. A number of relief agencies have clearly been able to successfully implement high-frequency surveys in regions as challenging as South Sudan, Darfur, Somalia, and northern Kenya (McKinney 2011). In most of these cases there was substantial technical leadership and oversight from international agencies. These surveys were often relatively small (or started small) and pragmatic in terms of data collection strategies. Surveying seminomadic pastoralist groups, for example, requires careful planning, though the widespread use of mobile phones and handheld computers in even the poorest countries now makes these kinds of surveys far more practical and cost-effective than in previous decades (Barnett and Gallegos 2013; Dillon 2012). We therefore conclude that domestic capacity is less of a long-term constraint than one might initially think.

TOWARD A NEW SYSTEM: A NETWORK OF HIGH-FREQUENCY, LONG-TERM SENTINEL SITES

The discussion above motivates our proposal for long-term, sustained collection of high-frequency, multidimensional data in strategically located sentinel sites targeted toward countries, regions, and agroecologies that show signs of persistent vulnerability to shocks and stressors. The primary target audiences are national governments, international donors, and implementing agencies that aspire to intervene—through emergency or nonemergency programs—in ways that reduce populations' susceptibility to episodes of severe acute malnutrition and food insecurity, enhance those populations' longer-term prospects for escaping and avoiding poverty, empower the most vulnerable to exert greater control over their own destiny, and safeguard the natural systems on which many of the poor's future livelihoods depend. The data generated by this system must provide timely, accurate inputs into decisions about resource mobilization and targeting, as well as help inform evaluation of interventions' impacts and updating of program designs. Secondary audiences would be the international community of researchers doing more basic research on related phenomena.

The data must be collected over the long term precisely because vulnerability to shocks emanates from slower-moving stressors, such as climate and environmental change, population growth, and sociopolitical transformation, and because the timing of shocks is, by definition, impossible to predict with certainty. These surveys therefore either need to be able to collect these kinds of data (through multidimensionality) or they need to be amenable to merging with such data (such as via georeferencing of survey sites). In any event, these processes cannot be understood and quantified with data collected over just a four- to five-year period. This must be a sustained effort.

The data must be collected at high frequency because assessing and understanding the impacts of shocks and external interventions requires recent preshock baselines. High-frequency data are also essential for

[14] This assessment is based on personal discussions with a number of individuals involved in these programs. One respondent also stated that the term *surveillance* had sometimes caused offense.

understanding the complex dynamics that underpin poverty traps and welfare dynamics more generally (Barrett and Carter 2013). What constitutes sufficiently high frequency is context specific. In HKI's NSP in Bangladesh, adequate funding and the complex high-rainfall (bimodal) climate in the country motivated HKI to collect data on a bimonthly basis. For similar reasons, Kenya's ALRMP has collected monthly data, although that has generally outstripped its technical capacity to collect, enter, clean, and analyze the data in a timely manner and thus seems an overreach. An FSMS in Malawi switched from surveying three times a year to twice a year because it was felt that the third survey was superfluous given the prevailing unimodal climate.[15]

The data must be multidimensional, covering food security, nutrition, and poverty, but also agricultural production and other economic activities, coping strategies, asset ownership, and community characteristics, at a bare minimum. Not all of these data need to be collected at high frequency, and in practice these surveys might consist of household/individual components and separate community surveys. Forethought should be given as to the potential to incorporate macroeconomic and ecological data from other sources.

Finally, we advocate for a very strategic collection of sentinel sites. A system of sentinel sites need not be nationally representative in the tradition of LSMS- or DHS-style surveys. Rather the sites can be targeted according to representativeness of different levels of vulnerability, different livelihoods, and different agroecologies. In this way, data collected at the individual and household levels can be harmonized with data collected at the community, macroeconomic, and environmental levels to facilitate a richer analysis of key drivers of change across time and space. In this context much can also be learned from existing programs, such as the Long Term Ecological Research (LTER) Network of 26 different sites that the US National Science Foundation created in 1980, with financial support from multiple other agencies, to provide a platform for rigorous field research on ecological issues that span vast spaces and unfold over many years or even decades. The LTER Network has generated unique datasets to study processes occurring over long time periods at multiple sites, leading to both timely early warning of unfolding ecological hazards and path-breaking research that has strongly informed environmental policy and the efficient allocation of limited resources. Not only have country-specific surveillance systems been successfully implemented already, but a network of sentinel sites has been operational for more than three decades with high impact in a different sphere of research and field-based programming (Hobbie et al. 2003). The adaptation to measurement and monitoring of resilience should be very feasible if the sites that constitute the network are chosen carefully.

[15] We thank Suresh Babu for this insight.

4. WHERE DO WE NEED TO INVEST IN A RESILIENCE MEASUREMENT SYSTEM?

Where should this kind of long-term, high-frequency measurement system be implemented? In this section we consider this issue but abstract from implementation issues, which we discuss in Section 6. We first focus on a very basic cross-country analysis to identify countries that are vulnerable according to different indicators of exposure to hazards and to child malnutrition and mortality. The greater costs associated with long-term, high-frequency surveys means that the number of sentinel sites will need to be limited and strategically targeted within vulnerable areas of a country, but across different types of livelihood zones carefully selected to characterize broader regions similar in agroecological and socioeconomic institutional terms. As an illustrative example, we zoom in on East Africa as a highly vulnerable region with substantial spatial diversity in underlying sources of vulnerability, and we review how sentinel sites might be chosen within that region.

IDENTIFYING THE MOST VULNERABLE COUNTRIES

While an international system of long-term, high-frequency sentinel sites for food and nutrition security measurement would ideally involve data collected in as large a number of developing countries as possible, resource constraints suggest that this system should first be developed in those countries that are most exposed and more vulnerable to natural—and to the extent feasible, man-made—hazards. To identify these countries we focus on indicators of several key dimensions of food and nutrition security:

1. **Natural disasters.** This indicator captures the total number of people affected by natural disasters over 1990–2013, relative to the average population over this period. *Affected* refers to people injured, killed, displaced, or in need of emergency assistance. This indicator captures both vulnerability and exposure to natural hazards, and is sourced from the EM-DAT International Emergency Disasters Database (OFDA and CRED 2013). It is the aggregate of various impacts of disasters, including deaths, homelessness, injuries, and other impacts.
2. **Emergency aid.** We take the average of per capita emergency and humanitarian aid receipts (in 2011 US dollars) over 2002–2011, the only years for which these data are available from OECD (2013). This indicator captures vulnerability to shocks through dependence on foreign aid to deal with these disasters. We fully acknowledge that this indicator may be biased by many factors, such as political biases or conflict, as well as the well-known biases in aid receipts for large countries. Nevertheless, since one important benefit of a sentinel system would be the improved targeting of international aid, this indicator seems relevant.
3. **Child wasting.** The prevalence of wasting in children younger than five, from the World Health Organization (WHO 2013), is a well-documented indicator of acute malnutrition. Acute malnutrition will not only reflect calorie deprivation, or "hunger," but may also reflect micronutrient deficiencies, poor sanitation and water supplies, and conceivably even genetic predispositions.
4. **Child stunting.** The prevalence of stunting in children younger than five, from WHO (2013), is the established measure of chronic malnutrition in children. Unlike wasting, stunting reflects the cumulative nutritional impacts of both shocks and stressors, so there is an argument to include this indicator in addition to wasting.
5. **Child mortality.** This is the rate of mortality (per 1,000) for children younger than five, sourced from the World Bank (2013). Although this is nominally a health indicator, almost half of child deaths are thought to be indirectly the result of poor nutrition (Black et al. 2013). This may therefore be a good catch-all indicator of poor nutrition, disease burdens, and the general quality of a country's health system, such as its ability to deliver vaccinations and to provide medical assistance before, during, and after birth.

The selection of these five indicators over a number of alternatives—or for that matter, a single indicator or index—is based on several justifications. First, data availability and data quality were primary concerns.[16] Second, these indicators are not particularly highly correlated, as Table 4.1 demonstrates. Emergency aid and natural disasters bear only modest (albeit statistically significant) correlations with other variables, in the 0.12–0.22 range, and even the child nutrition and health variables exhibit only moderately strong correlations. In other words, none of these indicators appears superfluous, and each could tell a different story if analyzed by itself. So the use of several different indicators lends robustness to our results.

Table 4.1 Bivariate correlation coefficients between the different indicators

	Emergency aid	Natural disasters	Child wasting	Child stunting
Natural disasters	0.19			
Child wasting	0.21	0.13		
Child stunting	0.13	0.22	0.40	
Child mortality	0.17	0.12	0.37	0.65

Source: Authors' estimates from data for 92 developing countries, from OECD (2013), OFDA and CRED (2013), WHO (2013), and World Bank (2013).
Note: *n* = 92 countries.

Table 4.2 (page 14) presents the raw data for all five indicators for the 30 most vulnerable countries, as defined by rankings of 1 through 30 for each indicator. For emergency aid per capita, the listing corresponds to countries receiving more than $2.50 per capita per year. For infant mortality, it approximately corresponds to countries in which 120 or more of every 1,000 children perish before age five. For natural disasters, it corresponds to 0.5 disasters per person over 1990–2013. For child stunting and child wasting, the listing approximately corresponds to those countries with prevalence rates above 40 percent for stunting and above 10 percent for wasting.

In addition to presenting the raw data and variable-specific rankings, we also categorize countries according to how many times they appear in the table, and we use this count to identify the highest levels of vulnerability. Countries in red appear four or five times (with one exception, discussed below). In other words, virtually all of these indicators signal that these countries are highly vulnerable and in great need of the kinds of measurement systems we advocate. Countries in yellow appear three times, suggesting that these countries are highly vulnerable in some of the measured dimensions but less vulnerable in others. Countries in gray appear just once or twice, suggesting that these are not the highest-priority countries for food and nutrition security sentinel sites.

In Table 4.3 (page 15) we extract the derived list of red (extremely vulnerable) and yellow (highly vulnerable) countries. The highest-priority list consists of 11 countries, all of which are in Africa south of the Sahara. This consists of four countries from the Horn of Africa: Ethiopia, Somalia, Sudan, Djibouti, and Eritrea. At the national level Kenya—also in the Horn—is a borderline case that only appears "highly vulnerable" from the dimensions of emergency aid, natural disasters, and child mortality, but large parts of pastoralist Kenya almost certainly belong to the "extremely vulnerable" category.[17] Chad, Mauritania, and Niger constitute a

[16] We were able to obtain these indicators for 92 low- and middle-income developing countries, including "fragile states" such as Somalia. Many alternative indicators of food and nutrition security were not measured for the full sample of countries, and the exclusion of countries is typically nonrandom. In some cases the quality of World Bank poverty or FAO food supply data for the most vulnerable countries may also be very poor, whereas the indicators above are—we believe—measured with substantially less error and greater international comparability.

[17] Subnational data from the Kenyan DHS show that wasting and infant mortality are much higher in these areas than in the Kenyan highlands. In the North Eastern region, for example, child wasting was 19.5 percent in the 2008–2009 round and 25 percent in the 2003 round. Moreover, Kenya appears fifth in the ranking of natural disaster exposure (Table 4.2), having had a very high incidence of drought in the last 15 years especially.

second, Sahelian cluster. As with Kenya, there are strong arguments to consider large parts of neighboring countries as being suitable for sentinel sites. DHS data tell us that child wasting is around 20 percent in Nigeria's North West and North East regions, while child mortality exceeds 200 per 1,000. There are also similarly high rates of wasting and mortality in Burkina Faso and Mali. Finally, Malawi, the Democratic Republic of Congo, and Burundi are a third set of African countries that also have very poor child health and nutrition indicators, and high exposure to shocks.

The next group of highly vulnerable countries (yellow) is larger and more diverse. It consists of 11 other countries of Africa south of the Sahara, covering southern, central, and western Africa. These countries are a mixed bag in terms of why they appear in this list. Mali and Burkina Faso constitute two more Sahelian countries that also have areas of extreme vulnerability to drought, but these countries are somewhat better off than Chad and Niger. Zambia is also exposed to drought, but seemingly somewhat less vulnerable than Malawi. Sierra Leone, Liberia, and Angola are resource-rich postconflict countries with very poor health and nutrition indicators, though no indications of high exposure to natural disasters. Uganda, Madagascar, and Mozambique finish out the African list of highly vulnerable countries, but are quite diverse.

Next, five South Asian countries appear on the list. All have unusually high rates of child wasting but also very high exposure to a range of shocks, including droughts, floods, storms, and earthquakes. As noted above, Bangladesh already has a nationwide sentinel survey in place. There are equally strong justifications for extending such a system to other parts of South Asia, particularly to vulnerable parts of India and Pakistan. Nepal, also, should arguably be on this list, although it suffers less from climatic extremes.

The remaining highly vulnerable countries are very diverse. Like the South Asian countries, Cambodia has high rates of wasting and stunting, and seems equally exposed and vulnerable to natural disasters. Haiti is the sole country in the Americas that appears to warrant a sentinel system. Yemen is the sole country in the Middle East and North Africa.

SELECTING SENTINEL SITES USING SPATIALLY DISAGGREGATED SUBNATIONAL DATA

Tremendous advances in the availability of spatially explicit data for the developing world enable us to select sentinel sites strategically, according to key dimensions of resilience. DHS data representative at the subnational level are available for most African countries and for many of the other highly vulnerable countries listed above. DHS provides useful indicators of child wasting, stunting, and mortality rates, as well as maternal body mass index. But within countries and regions with alarmingly high indicators it will be important to strategically select sentinel sites across different agroecological and livelihood zones so as to ensure cost-effectiveness in representing the diverse peoples of these regions. In this regard there are several options.

In recent years there have been a number of efforts to better map livelihood zones, largely for the purposes of strategically directing development and relief efforts. This effort began with the "development domains" literature, which used geographic information systems data on agroecological potential, market access, and population density (Chamberlin, Pender, and Yu 2006). The Harvest Choice initiative at the International Food Policy Research Institute (IFPRI) has since expanded the measurement of these kinds of variables. And more recently USAID has funded livelihood mapping through the use of focus group surveys in which respondents describe the salient features of the local community. The best such example is the rich *Atlas of Ethiopian Livelihoods* (USAID and Ethiopia, DRMFSS 2011).

Table 4.2 Country rankings and scores on five different indicators of exposure and vulnerability to shocks

Country	Emergency aid per capita (US$)	Country	Infant mortality	Country	Natural disasters	Country	Under-5 stunting	Country	Under-5 wasting
Somalia	16.01	Sierra Leone	273	China	2.18	Burundi	63.1	Djibouti	25.7
Haiti	14.14	Angola	260	Somalia	1.83	Afghanistan	59.3	Sudan	21.0
Liberia	13.21	Niger	260	Malawi	1.77	Niger	54.5	Burkina Faso	18.7
Sudan	9.65	Liberia	235	Niger	1.67	Guatemala	54.3	India	17.1
Afghanistan	9.33	Mali	220	Kenya	1.65	Ethiopia	54.1	Bangladesh	16.2
Djibouti	7.71	Chad	207	Djibouti	1.52	Malawi	53.6	Indonesia	14.8
Eritrea	6.66	DRC	205	Philippines	1.49	Nepal	53.2	Sri Lanka	14.7
Chad	6.40	Burkina Faso	200	Cambodia	1.37	Yemen	53.1	Pakistan	14.2
Burundi	5.73	Nigeria	197	Eritrea	1.34	Madagascar	52.8	Mali	13.7
Zimbabwe	5.30	Zambia	182	Mongolia	1.31	Bangladesh	51.6	Chad	13.5
Mauritania	5.23	Burundi	181	Thailand	1.24	Angola	50.8	Somalia	13.2
Sierra Leone	4.69	CAR	179	Zimbabwe	1.22	Rwanda	50.0	Madagascar	13.0
Ethiopia	4.09	Rwanda	169	Mauritania	1.13	Lao PDR	48.2	Yemen	12.9
Nicaragua	3.99	Mozambique	154	Namibia	1.13	Sudan	47.6	Mauritania	12.8
Niger	3.69	Somalia	153	Bangladesh	1.12	Mozambique	47.0	Nigeria	12.5
Mongolia	3.58	Benin	153	Tajikistan	1.05	Cambodia	46.5	Eritrea	12.4
Sri Lanka	3.53	Cameroon	150	Haiti	0.93	DRC	45.1	Togo	12.3
Guyana	3.34	Uganda	138	Mozambique	0.90	CAR	44.6	Nepal	11.8
CAR	3.26	Djibouti	137	India	0.89	Tanzania	44.4	Uzbekistan	11.6
Jordan	3.15	Malawi	134	Lao PDR	0.89	PNG	43.9	Syria	11.5
Gambia	2.88	Ethiopia	133	Zambia	0.77	Eritrea	43.7	Cambodia	10.8
Bolivia	2.83	Côte d'Ivoire	131	Chad	0.73	Zambia	43.0	Myanmar	10.7
Honduras	2.75	Tanzania	127	Sri Lanka	0.71	Sierra Leone	42.7	Niger	10.7
Malawi	2.71	Mauritania	125	Ethiopia	0.68	Liberia	42.4	Côte d'Ivoire	10.6
Uganda	2.66	Madagascar	124	Honduras	0.64	Nigeria	42.1	Ethiopia	10.5
Angola	2.59	Congo, Rep. of	123	Sudan	0.63	Somalia	42.1	Guinea	9.8
Kenya	2.59	Senegal	122	Burundi	0.62	Chad	42.1	Vietnam	9.7
DRC	2.56	Gambia	121	Peru	0.52	Benin	41.9	DRC	9.3
Tajikistan	2.56	Kenya	119	Mali	0.52	Uganda	41.8	Morocco	9.3
Guatemala	2.52	Ghana	117	Nicaragua	0.52	Pakistan	41.5	Haiti	9.2

Source: Authors' construction from OECD (2013), OFDA and CRED (2013), WHO (2013), and World Bank (2013).

Notes: CAR = Central African Republic; DRC = Democratic Republic of Congo; Lao PDR = Lao People's Democratic Republic, PNG = Papua New Guinea.

Legend:

Highly vulnerable	Moderately vulnerable	Less vulnerable

Table 4.3 The most vulnerable countries according to the five indicators

Extremely vulnerable	Highly vulnerable	
Burundi	Angola	
Chad	Central African Republic	Bangladesh
Democratic Republic of Congo	Kenya	Cambodia
Djibouti	Liberia	Sri Lanka
Eritrea	Madagascar	
Ethiopia	Mali	Haiti
Malawi	Mozambique	
Mauritania	Nigeria	Yemen
Niger	Sierra Leone	
Somalia	Uganda	
Sudan	Zambia	

Source: Authors' construction from OECD (2013), OFDA and CRED (2013), WHO (2013), and World Bank (2013).

To give a sense of how one can go about this, we present a brief example from Ethiopia in Figure 4.1. This is a pertinent example given Ethiopia's high levels of vulnerability (see above), its tremendous spatial diversity, and the existence of a number of large-scale programs designed to build household and community resilience. How, then, would one go about selecting sentinel sites in such a country? First, we have created a layer of agroecological data to define five traditional agroecological zones in Ethiopia: pastoralist lowlands, drought-prone highlands (with a mix of agrarian and agropastoralist livelihoods), rainfall-sufficient highlands focused around *enset* (false banana)–based production systems, rainfall-sufficient highlands with a focus on cereals, and humid lowlands confined to the west of the country. Second, we separately add on two other layers of village (cluster)–level estimates of wasting prevalence (Panel A) and stunting prevalence (Panel B) from the Ethiopian DHS of 2011.[18] Darker colors represent higher levels of wasting/stunting.[19] Wasting generally captures acute (short-term) malnutrition that could be related to food deprivation or disease burdens. Stunting is a more cumulative, longer-term measure that represents a much broader range of factors over a child's life cycle, especially the first 1,000 days of life—in the womb and the first two years of childhood.

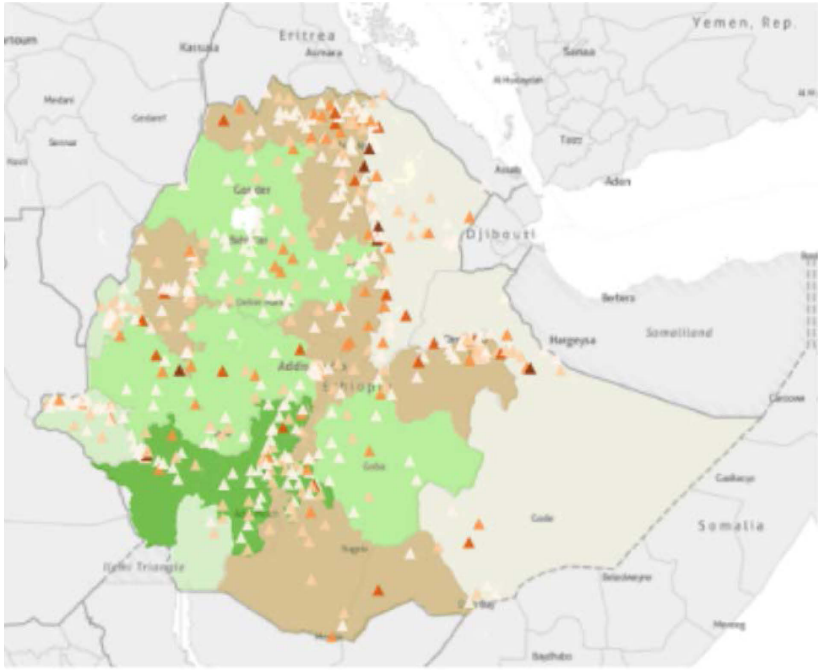
It is evident from Panel A that there are clear patterns of wasting in specific agroecological zones. Though undersurveyed, wasting is clearly higher in drier and drought-prone pastoralist areas, and averages a very high 20 percent in the Afar and Somali regions. Other drought-prone highland areas are also vulnerable. Hence it certainly appears that vulnerability to shocks is indeed much higher in these areas. Panel B tells a somewhat different story. Stunting prevalence is more evenly spread through the country and not confined to particular regions. Thus if the focus of the sentinel system is on monitoring acute malnutrition, then this suggests that sentinel sites should be targeted at drought-prone areas in both the highlands and the lowlands, along with control group villages in relatively similar agroecologies (in order to understand why some communities appear vulnerable while otherwise similar communities appear resilient).

[18] Note that DHS surveys are not representative at the cluster level. Instead our goal here is just to determine basic spatial patterns rather than to make specific statistical inferences.

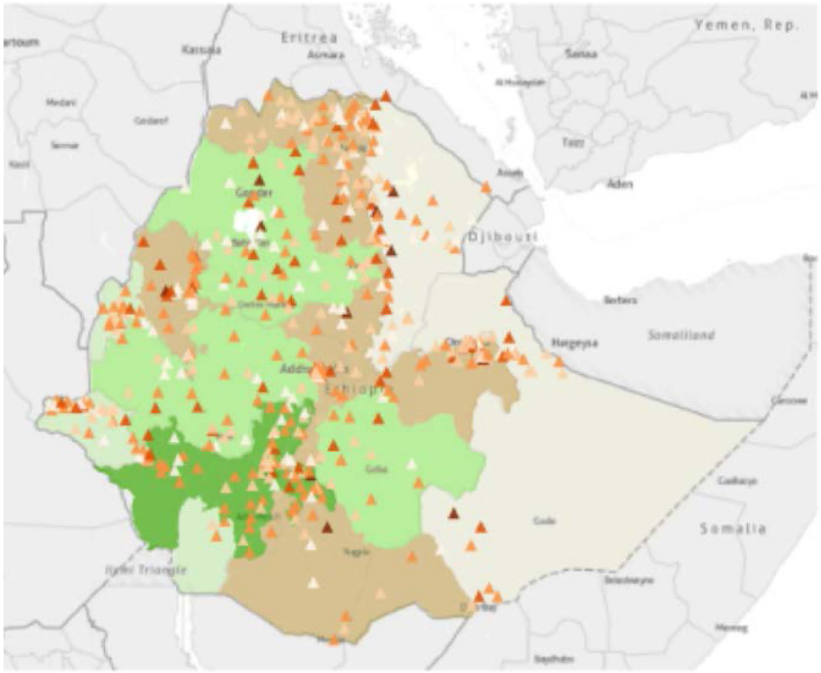
[19] Naturally, other indicators could be used here, including maternal wasting, child mortality, vulnerability to droughts, market access, and so on.

Figure 4.1 Village-level wasting and stunting by agroecological zone

Panel A: Wasting and agroecological zones in Ethiopia, 2011



Panel B: Stunting and agroecological zones in Ethiopia, 2011



Children Moderately Wasted, %

- 0-10
- 10-20
- 20-30
- 30-40
- > 40

Agro-Ecological Zones

- pastoralist lowlands
- drought-prone highlands
- rainfall-sufficient highlands - cereals
- rainfall-sufficient highlands - enset
- humid lowlands

Children Moderately Stunted, %

- 0-20
- 20-40
- 40-60
- 60-80
- 80-100

Agro-Ecological Zones

- pastoralist lowlands
- drought-prone highlands
- rainfall-sufficient highlands - cereals
- rainfall-sufficient highlands - enset
- humid lowlands

Source: Maps provided by Carlo Azzarri and Melanie Bacou of the International Food Policy Research Institute (IFPRI), based on data from the Ethiopian Demographic and Health Survey of 2011.

Figure 4.1 is merely an illustrative example from one country. But it underscores two key points. First, existing data sources are sufficient to identify extremely and highly vulnerable areas by livelihood zone in ways that can usefully inform the location of sentinel sites. Second, agroecological characteristics are strongly correlated with food-security indicators based on child health and nutritional status. Georeferenced data are becoming more available and on a more timely basis, especially as information and communications technologies (ICTs) enable increased crowdsourcing of data. We have the data to focus attention on high-frequency, high-quality, long-term data collection in the places most vulnerable to shocks and widespread acute human suffering.

5. HOW SHOULD WE MEASURE RESILIENCE? DATA COLLECTION DESIGN

Once sentinel survey sites have been strategically selected, the next question is what data to collect, how, and from whom—that is, the sampling strategy and appropriate data collection instruments. For any sustained, long-term effort, a common baseline survey repeated at regular intervals is essential, perhaps supplemented by occasional special-purpose survey modules, experiments, and other data collection instruments to inform specific questions. The core survey instrument must provide reliable, timely measures that meet the central characteristics of the emergent definition of resilience in development and relief applications, discussed in Section 2. The need to address dynamics, stressors, shocks, and systems carries particular implications for the design that we explore in this section. The key questions are these: What is the minimum necessary frequency of surveys for resilience measurement? Should these surveys be panels (that is, repeated observations of the same individuals, households, or villages) or simply repeated cross-sections? And what sorts of variables should these surveys measure? After discussing the salient trade-offs related to these issues, we outline a relatively novel hybrid strategy for achieving some of the multiple goals of resilience measurement.

HOW FREQUENTLY SHOULD DATA BE COLLECTED?

There is clearly a compelling case for high-frequency measurement, but the higher the frequency of measurement the greater the cost and the higher the risk of measurement errors induced by rushed data collection, entry, and cleaning. Moreover, in terms of the total budget, there will clearly be trade-offs between spatial coverage (the number of survey sites) and temporal coverage (the frequency of surveys within sites).

Experience offers some useful guidance here. HKI's original NSP in Bangladesh ran surveys on a bimonthly basis, nominally on the grounds that Bangladesh has complex rainfall patterns. The government of Kenya's ALRMP data collection has been monthly. But both are perhaps unusual cases of very well-funded programs. And the ALRMP data management (postcollection) system has often been overwhelmed by the flow of data, sometimes leading to quality control issues. At the other extreme, the Food Security and Nutrition Monitoring program conducted in Malawi in the 1990s was implemented twice a year to coincide with the agricultural production cycle and the preharvest hunger season (Babu and Mthindi 1994). In that survey the implementing team initially ran three rounds per year but concluded that one of the rounds was superfluous and dropped to twice annually.

Semiannual surveys would reduce cost, but they would also come at the expense of loss of information content, including reducing the efficacy of these surveys as an early-warning system. One option we discuss below is conducting "thick rounds" semiannually (to collect a broader range of data, including data on agricultural production) but conducting "thin" rounds more frequently (to focus solely on food and nutrition indicators and on questions that might be useful for early-warning purposes).

REPEATED CROSS-SECTIONS OR PANELS?

Much economic research on poverty dynamics and vulnerability exploits panel data, a practice that has the obvious advantage of allowing analysts to track individuals or individual households over time in order to study the impacts of idiosyncratic shocks (such as illness) as well as the idiosyncratic impacts of covariate shocks, including the efficacy of different individual- or household-level coping strategies. Panel data are also advantageous for large-scale program evaluation, since they offer much greater scope to identify causal impacts of policies and programs (Elbers and Gunning 2013). Repeated cross-sections would be useful in this regard only if there were significant spatial variation in program implementation.

Repeated cross-sections, on the other hand, have two advantages. First, nutritional surveillance has often used repeated cross-sections in order to focus on young children and pregnant mothers. Thus the NSP in Bangladesh focused on surveying households that had at least one young child. On the other hand, this apparent advantage is a slim one in high-fertility settings in which most households have young children

present. Second, panel surveys are more expensive and more challenging in contexts where it is difficult to track respondents down. Seminomadic populations are a case in point, although mobile telephones increasingly offer a solution to the challenge of tracking survey respondents.

Again, we note that hybrid approaches are also possible. The survey could be principally designed as a panel but could contain a nonpanel component to oversample a specific group. In particular, the international community's reasonable concern for early childhood health and nutrition indicators would justify supplementing a core panel survey with repeated cross-section sampling to ensure statistically representative coverage of targeted subpopulations of particular interest, particularly pregnant and lactating women, and children younger than five.

WHAT NEEDS TO BE MEASURED?

Existing surveys differ greatly in terms of what they seek to measure and how they measure it. Table 5.1 provides an overview of the key domains of interest, drawing on Barrett and Constas (2013) and FSIN (2013). In terms of core welfare indicators, the nutrition literature has well-established norms for measuring acute malnutrition (weight for height, mid-upper-arm circumference, body mass indexes, and even hemoglobin levels), but there is far more debate on food-security measurement (Barrett 2010; Headey and Ecker 2013; Headey, Ecker, and Trinh Tan 2014). In high-frequency surveys, detailed data on household expenditure, income, and calorie consumption will typically be too time-consuming and expensive to collect. Alternatives include dietary diversity/quality indexes, such as WFP's Food Consumption Score (FCS), and more subjective indicators, such as the Household Food Insecurity and Access Score (HFIAS). While there are advocates of both types of indicators, recent findings favor dietary diversity scores on the grounds that they are sensitive to shocks (Headey, Ecker, and Trinh Tan 2014); are a good proxy for calorie consumption (Wiesmann et al. 2009); and are nutrition sensitive, especially when collected for young children and mothers (Headey and Ecker 2013; Tiwari, Skoufias, and Sherpa 2013). So we would advocate for FCS measures in both the "thick" and "thin" rounds, perhaps supplemented with HFIAS-style subjective indicators in selected "thick" rounds or supplemental modules. Health indicators related to disease, injury, and disabilities will also be important, given the well-documented interactions between food insecurity and health.

In addition to these core welfare indicators, it is very important to measure variables that tell us about the underlying sources of vulnerability in a population, including stressors. These would include a variety of household and community assets; demographic factors; and broader economic, social, and ecological factors. Some of those data can be gathered at community scale. Some can be collected from external (for example, satellite-based remote sensing) platforms and integrated with the survey data collected within the community. Indeed, suitably processed externally sourced data can and should be delivered directly to sentinel communities so as to help inform their own actions, not just those of external actors. Subjective assessments of the relevance and relative severity of different stressors can be elicited using now-well-established methods (Manski 2004; Delavande, Giné, and McKenzie 2011), but the empirical evidence thus far suggests that most variation in such subjective indicators is spatial and intertemporal, not among households within a community during a given period (Doss, McPeak, and Barrett 2008).

Next come the shocks. Some shocks can be measured at more aggregate levels (for example, rainfall, or vegetation indexes), but previous research has found that covariate shocks typically have very different impacts across households even within local communities. Thus, in addition to measuring shocks according to changes in core welfare indicators and assets, subjective questions on both covariate and idiosyncratic shocks will be important. Similarly, there is now a well-established approach to measuring coping strategies—in particular the coping strategies index (Maxwell and Caldwell 2008)—although there will often be a great deal of context specificity in this method, and there is a need in longer-term surveys to more carefully measure adaptation. Finally, for the purposes of monitoring and evaluation, it will be critical to ask detailed questions about access to relief programs.

Table 5.1 Proposed metrics for measuring resilience to food and nutrition insecurity

State variables	Shocks, stressors, and responses
Core welfare indicators • Food security • Nutrition outcomes • Health outcomes *Assets and other initial conditions* • Demography • Human capital • Physical assets • Livelihoods • Social and political capital • Access to services • Access to infrastructure • Mitigation strategies • Food, fuel, and other prices *Economic, social, and ecological environment* • Ecology and environment • Policies and institutions	*Covariate (large-scale) shocks* • Drought/flood • Health shocks (such as epidemics) • Political crises • Economic shocks • Agronomic shocks *Idiosyncratic (household-specific) shocks* • Illness/death • Loss of income • Crop loss • Livestock loss *Responses* • Household coping/adaptation strategies • Community coping/adaptation strategies • External relief efforts

Source: Authors' adaptation from Constas and Barrett (2013).

TOWARD A HYBRID APPROACH

We emphasize again the critical importance of starting from the actions that (raw or processed) data might inform and working backward to select sites, determine sampling strategies, and guide survey design. For a new and more ambitious system of sentinel sites to be cost-effective in informing feasible actions, the network would need to be built around data collection systems that (1) offer a wide range of benefits for governments, relief agencies, development agencies, and researchers, and (2) keep costs down as much as possible, without forsaking major benefits. This will be a delicate balance, and there are grounds to consider how various innovations in design and technology use could help achieve the right balance.

In that vein we make several suggestions. First, we recommend mixing together so-called thick and thin rounds. Thick rounds could be conducted less frequently (for example, semiannually, annually) but would involve collection of the widest possible array of data. Thin rounds would instead focus on variables that vary markedly over time, such as food and nutrition security indicators (including potential early-warning indicators); subjective indicators of stressors, shocks, and coping strategies; and access to relief services. As a rule, all thin-round data would also be collected during thick rounds. Some thin rounds could, however, be reserved for special-purpose, one-time surveys.

Second, at least after initial familiarization of both respondents and survey teams, thin rounds could potentially be implemented remotely through the use of mobile phones or other ICTs. These approaches would, of course, need to be subjected to validation exercises, and there are clearly circumstances—such as when a severe slow-onset disaster, like a drought, is obviously emerging—when ICTs might be appropriately replaced by traditional methods, and when thin rounds should be "thickened" with additional questions or additional sampling. But it seems probable that with additional experimentation and learning-by-doing, creative use of ICTs could prove effective for shorter surveys and especially for mobile populations. There is currently substantial experimentation with ICT-based surveys in a number of countries, and in the next couple of years we will know much more about exactly what can and cannot be achieved with these approaches (Barnett and Gallegos 2013; Bauer et al. 2013; Dillon 2012).[20]

[20] See also http://www.humanitarianinnovation.org/blog/1442.

Third, as was argued above, this sentinel system should focus on panel data in order to capture the stochastic dynamics of the key outcome indicators of interest, but also oversample and ensure statistical representativeness of demographic groups of particular interest, such as pregnant and lactating women.

Fourth, a sizable proportion of the most vulnerable communities in the developing world are seminomadic, due to seasonal migration for either employment or herding, and flexible approaches will be needed for such communities. For the most part these groups follow reasonably structured migration routes, returning to the same sites at specific times of the year. These kinds of surveys will therefore need to be timed according to these migration patterns, perhaps surveying the same community in one location at one time of the year and at other locations at other times of the year. The World Bank, IFPRI, ILRI-Cornell and USAID are already conducting panel surveys of pastoralist groups in Ethiopia and Kenya, and as we noted above, the Arid Lands project has been doing this in Kenya for many years. Much can be gained by learning from these experiences. And we again emphasize that mobile telephony and other ICTs (for example, global positioning systems) now make this kind of flexible survey arrangement much more feasible and cost-effective than it would have been in the past, though further experimentation is still needed.

Fifth, with the dramatic growth in attention paid to rigorous impact evaluation of development and relief interventions, the sentinel system would offer an ideal platform for experimental and quasi-experimental evaluations of the efficacy and effectiveness of alternative approaches to building resilience. Combined with rich, long-term observational data from the panel surveys, RCTs and other experimental designs could be embedded to generate short-run evaluation results, while the ongoing panel monitoring would permit longer-term assessment of the persistence of impacts postintervention. Since a key undercurrent of the growing enthusiasm for resilience as an organizing concept behind development and relief efforts among chronically vulnerable populations is the aspiration to achieve lasting impacts from short- to medium-term interventions, the sort of sentinel system we propose offers a *necessary* platform, but one presently lacking in existing approaches.

Finally, we emphasize that all these recommendations should be the subject of experimentation, refinement and ongoing learning. Whilst the international community must act boldly in pushing this initiative ahead, the complexity of developing a system of sentinel sites in the world's most vulnerable and fragile regions clearly necessitates an ongoing process of "learning by doing." This could entail starting with pilot projects in each of the relevant countries or subnational regions to first improve and refine the survey instruments and data collection processes, before scaling up. It would also entail cross-country learning through appropriate institutional platforms (more on this below). Yet previous experience in countries as diverse as Bangladesh and Kenya shows that these kinds of surveys can be done even in the most challenging environments.

6. MOVING FORWARD

While many of the arguments and suggestions made in this paper would apply to individual surveys, we have explicitly proposed a more ambitious and far-reaching approach that would involve the establishment of a global data collection network of long-term, high-frequency sentinel sites in the world's most vulnerable regions. This would, at the very least, include countries that are perennial, large-scale recipients of humanitarian assistance and regularly suffer major covariate shocks such as droughts, floods, storms, or war with an unacceptable baseline level of human health, nutrition, and poverty indicators. The cost of these surveys would be substantial, although we perceive the likelihood of far greater benefits than costs, provided that the design is rooted from the beginning in demand from different government agencies, bilateral and multilateral donors, operational agencies, and research communities. That level of demand will require paying special care to strategies for high-quality data collection, cleaning, management, and dissemination, as well as data processing and investments in analytical capacity to spread the knowledge benefits widely and capture these varied communities' full willingness to pay for this valuable public good.

Such a system would essentially be new, rather than an extension of existing systems. This does not mean reinventing the wheel—as we discussed in Section 3, there are existing survey approaches that have done a very effective job of improving our measurement and understanding of resilience—but an essentially new system is required for several important reasons. First, few existing approaches are conducted at scale across multiple countries with long-term time horizons. Second, many existing approaches have a strong unidisciplinary rather than a multidisciplinary focus (for example, *nutrition* surveillance systems). And third, this expansive scale and multidisciplinary scope will certainly require new institutional and financial arrangements.

These arrangements should consist of a consortium of international institutions, which would then work in partnership with national governments and NGOs. The international consortium should consist of a mix of institutions that are consumers of the information generated and others that can contribute to the rigorous design or implementation of data collection, dissemination, and processing. A nonexhaustive list of key stakeholders would include the following:

- **National governments.** Country-level cooperation and "ownership" of these surveys will be important, though some governments will need persuasion and some of these countries (or regions within these countries) are so-called failed states where government is effectively absent or counterproductive. Indeed, the lack of state capacity is a commonly cited explanation of why some of these regions are vulnerable in the first place. Nevertheless, it is indeed crucial to obtain national support and search for champions within national governments. Moreover, an explicit objective of this survey system should be, over time, to build national capacity to conduct and analyze high-quality survey data as part of broader resilience-building strategies.
- **Bilateral donors and foundations.** The support of major bilateral donors and their key agencies would be essential for demonstrating and financing long-term commitments for these surveys. There is substantial precedent in this regard. USAID's funding of DHS and the Famine Early Warning System Network (FEWSNet) has surely been among its higher-return investments; DHS has been truly revolutionary for the monitoring and analysis of nutrition, health, demography, and several other fields, while FEWSNet has enabled considerably faster and more systematic response to slow-onset disasters. The sentinel system should also be able to provide a valuable research platform of the sort that the US government invested in when its National Science Foundation launched the LTER Network in 1980. The Bill & Melinda Gates Foundation has similarly funded major new statistical systems such as Harvest Choice and the Living Standards Measurement Surveys–Integrated Surveys on Agriculture (LSMS-ISA) projects for improving agricultural statistics. An equally ambitious strategy for measuring vulnerability and resilience in the countries that need it most requires a similarly high level of support from bilateral agencies.

- **Multilateral organizations.** As part of its Vulnerability Analysis and Mapping activities, the WFP already collects a significant amount of data on food and nutrition security. It has a permanent presence in the world's most vulnerable countries, and substantial technical capacity and experience to help implement these surveys. In some sense, the WFP has already embarked down the right road, but its progress is apparently constrained by finances; by a narrower (but understandable) interest in a few key indicators; and by the limits of its technical expertise in agriculture, ecology, economics, and other relevant scientific fields. UNICEF also has a mandate to monitor, measure, and address maternal and child nutrition, as does the WHO. Both could bring substantial technical expertise to the table, as could other UN agencies such as the UN High Commissioner for Refugees, UN Development Programme, UN Environment Programme, UN Human Settlements Programme (UN-HABITAT), and UN Population Fund. The World Bank has considerable technical expertise in large-scale surveys and in data management and dissemination, especially to government policymakers. For example, the World Bank's current LSMS-ISA project has substantially improved agricultural survey measurement techniques. The Food and Agriculture Organization of the UN (FAO) works closely with WFP on food-security assessments and runs the Global Information and Early Warning System, which offers larger-scale secondary data that would be valuable to integrate with higher-resolution, survey-based data from sentinel sites as well as with electronic dissemination platforms.
- **Major international and local NGOs.** Large international NGOs that have a presence in multiple countries clearly have a demand for better measurement of vulnerability and resilience, but many also have a permanent presence in these countries and substantial capacity to help implement these surveys. Again, there are already platforms for increased coordination among NGOs and various international agencies. A nonexhaustive list would include the International Red Cross and Red Crescent Societies, Action Against Hunger (ACF International), BRAC, CARE, Catholic Relief Services, Concern Worldwide, Médicins sans Frontières, Oxfam, World Vision, and of course, HKI, given its substantial experience on nutrition surveillance systems in Bangladesh and Indonesia. Major local NGOs would be natural partners in specific sites given their local expertise and influence.
- **CGIAR.** The network of international agricultural research centers that make up CGIAR could contribute technical expertise on survey design and structure as well as quality control, especially on agricultural and nutrition-focused surveys. CGIAR has a distinguished history in improved global food security, strong partnerships with national and regional organizations, and considerable expertise in project and program evaluation.
- **Leading international academic institutions.** A wide variety of experts from academia would be needed to supplement the technical expertise of the institutions listed above.

The challenge of coordination among so many partners is admittedly considerable. But this is also an important means of ensuring widespread use of data for both monitoring and research purposes, including program evaluation. Moreover, the coordination challenge is less than first meets the eye, because there is a clear sequence to be followed. The first step toward achieving the goal of a multicountry system of high-frequency, long-term sentinel surveys will be the creation of a consortium of donors. This consortium needs to make long-term commitments, albeit with provisos to assess the efficacy of the system at regular intervals. Then a management structure will need to be decided upon, along with country-specific management structures that reflect local relationships and resource availability. Next, an overarching technical steering group and various country-specific technical working groups will need to be established, drawing upon the staff of UN agencies, the World Bank, CGIAR, universities, and private consultancies. A central purpose of the technical steering group will be to ensure ongoing learning from field data collection experiences in each of the sentinel sites and in parallel efforts outside of the sentinel system, as well as from user group experiences with processed data and analysis. But the greatest step in this process will be the up-front commitment of international donors and national governments to revolutionize our collective ability to monitor and diagnose the problems that poor and vulnerable populations face on a perennial basis.

REFERENCES

Akhter, N., and N. Haselow. 2010. "Using Data from a Nationally Representative Nutrition Surveillance System to Assess Trends and Influence Nutrition Programs and Policy." *Field Actions Science Reports* 4. http://factsreports.revues.org/395.

Almond, D., and J. Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25 (3): 153–172.

Babu, S. C., and G. B. Mthindi. 1994. "Household Food Security and Nutrition Monitoring: The Malawi Approach to Development Planning and Policy Interventions." *Food Policy* 19:272–284.

Babu, S. C., and V. J. Quinn. 1994. "Food Security and Nutrition Monitoring in Africa: Introduction and Historical Background." *Food Policy* 19:211–217.

Barnett, I., and J. V. Gallegos. 2013. *Using Mobile Phones for Nutrition Surveillance: A Review of Evidence*. Transform Nutrition Evidence Paper 1. Sussex, UK: Institute for Development Studies.

Barrett, C. B. 2010. "Measuring Food Insecurity." *Science* 327:825–828.

Barrett, C. B., and M. R. Carter. 2013. "The Economics of Poverty Traps and Persistent Poverty: Policy and Empirical Implications." *Journal of Development Studies* 49 (7): 976–990.

Barrett, C. B., and M. Constas. 2013. "Toward a Theory of Resilience for International Development Applications." Unpublished working paper, Cornell University, Ithaca, NY, US.

Barrett, C. B., and B. M. Swallow. 2006. "Fractal Poverty Traps." *World Development* 34 (1): 1–15.

Bauer, J.-M., K. Akakpo, M. Enlund, and S. Passeri. 2013. "A Bridge Too Far? Mobile Text Food Security Surveys in North Kivu, Eastern Democratic Republic of Congo." Unpublished working paper, United Nations World Food Programme, Rome.

Béné, C., R. Godfrey-Wood, A. Newsham, and M. Davies. 2012. *Resilience: New Utopia or New Tyranny? Reflection about the Potentials and Limits of the Concept of Resilience in Relation to Vulnerability Reduction Programmes*. Working Paper 405. Brighton, UK: Institute of Development Studies.

Béné, C., D. Headey, L. Haddad, and K. von Grebmer. 2013. "Understanding Resilience for Food and Nutrition Security." In *2013 Global Hunger Index—The Challenge of Hunger: Building Resilience to Achieve Food and Nutrition Security,* 19–31. Bonn: Deutsche Welthungerhilfe (German AgroAction); Washington, DC: International Food Policy Research Institute; Dublin: Concern Worldwide.

Black, R. E., C. G. Victora, S. P. Walker, Z. A. Bhutta, P. Christian, M. de Onis, M. Ezzati, S. Grantham-McGregor, J. Katz, R. Martorell, and R. Uauy. 2013. "Maternal and Child Undernutrition and Overweight in Low-Income and Middle-Income Countries." *The Lancet* 382:427–451.

Block, S., L. Kiess, P. Webb, S. Kosen, R. Moench-Pfanner, M. W. Bloem, and C. P. Timmer. 2004. "Macro Shocks and Micro Outcomes: Child Nutrition during Indonesia's Crisis." *Economics and Human Biology* 2:21–44.

Bloem, M. W., R. Moench-Pfanner, and D. Panagides. 2003. *Health and Nutritional Surveillance for Development*. Singapore: Helen Keller Worldwide.

Carter, M. R., and C. B. Barrett. 2006. "The Economics of Poverty Traps and Persistent Poverty: An Asset-Based Approach." *Journal of Development Studies* 42 (2): 178–199.

Chamberlin, J., J. Pender, and B. Yu. 2006. *Development Domains for Ethiopia: Capturing the Geographical Context of Smallholder Development Options*. ESSP Working Paper 6. Addis Ababa: Ethiopia Strategy Support Program, International Food Policy Research Institute.

Chantarat, S., A. G. Mude, C. B. Barrett, and M. R. Carter. 2013. "Designing Index-Based Livestock Insurance for Managing Asset Risk in Northern Kenya." *Journal of Risk and Insurance* 80 (1): 205–237.

Constas, M., and C. B. Barrett. 2013. "Principles of Resilience Measurement for Food Insecurity: Metrics, Mechanisms, and Implementation Plans." Paper presented at Expert Consultation on Resilience Measurement Related to Food Security sponsored by Food and Agricultural Organization and World Food Programme, Rome, February 19–21.

Delavande, A., X. Giné, and D. McKenzie. 2011. "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence." *Journal of Development Economics* 94 (2): 151–163.

Devereux, S. 2006. *The New Famines: Why Famines Persist in an Era of Globalization*. London: Routledge.

Dillon, B. 2012. "Using Mobile Phones to Collect Panel Data in Developing Countries." *Journal of International Development* 24:518–527.

Doss, C., J. McPeak, and C. B. Barrett. 2008. "Interpersonal, Intertemporal, and Spatial Variation in Risk Perceptions: Evidence from East Africa." *World Development* 36 (8): 1453–1468.

Elbers, C., and J. W. Gunning. 2013. "Evaluation of Development Programs: Randomized Controlled Trials or Regressions?" *The World Bank Economic Review*. Published electronically September 7. doi:10.1093/wber/lht025.

FAO (Food and Agriculture Organization of the United Nations). 1996. *Rome Declaration on World Food Security and World Food Summit Plan of Action.* Rome. www.fao.org/docrep/003/w3613e/w3613e00.HTM.

FSIN (Food Security Information Network). 2013. *Resilience Measurement Principles: Toward an Agenda for Measurement Design.* Resilience Measurement Technical Working Group Technical Series 1. Rome: World Food Programme.

Griliches, Z., and J. A. Hausman. 1986. "Errors in Variables in Panel Data." *Journal of Econometrics* 31 (1): 93–118.

Headey, D., and O. Ecker. 2013. "Rethinking the Measurement of Food Security: From First Principles to Best Practice." *Food Security* 5:327–343.

Headey, D., O. Ecker, and J.-F. Trinh Tan. 2014. "Shocks to the System: Measuring Food Security in a Volatile World." In *Handbook on Food: Demand, Supply, Sustainability and Security*, edited by R. Jha, R. Gaiha, and A. B. Deolalikar, 41–71. Cheltenham, UK: Edward Elgar.

Headey, D., A. S. Taffesse, and L. You. 2014. "Diversification and Development in Pastoralist Ethiopia." *World Development* 56:200–213.

Hobbie, J. E., S. R. Carpenter, N. B. Grimm, J. R. Gosz, and T. R. Seastedt. 2003. "The US Long Term Ecological Research Program." *BioScience* 53 (1): 21–32.

Hoddinott, J., J. A. Maluccio, J. R. Behrman, R. Flores, and R. Martorell. 2008. "Effect of a Nutrition Intervention during Early Childhood on Economic Productivity in Guatemalan Adults." *The Lancet* 371 (9610): 411–416.

IPCC (Intergovernmental Panel on Climate Change). 2012. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*. New York: Cambridge University Press.

Johnson, N., and A. Wambile, eds. 2011. *The Impacts of the Arid Lands Resource Management Project (ALRMPII) on Livelihoods and Vulnerability in the Arid and Semi-arid Lands of Kenya.* Research Report 25. Nairobi: International Livestock Research Institute.

Jonsson, U. 1995. "Towards an Improved Strategy for Nutrition Surveillance." *Food and Nutrition Bulletin* 16 (2): 102–111.

Manski, C. F. 2004. "Measuring Expectations." *Econometrica* 72 (5): 1329–1376.

Maxwell, D., and R. Caldwell. 2008. *The Coping Strategies Index: Field Methods Manual.* Atlanta, GA, US: CARE. http://home.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp211058.pdf.

McKinney, P. 2011. *Module 10: Nutrition Information and Surveillance Systems, The Harmonised Training Package: Resource Material for Training on Nutrition in Emergencies*. Geneva: UNICEF.

Mude, A. G., C. B. Barrett, J. G. McPeak, R. Kaitho, and P. Kristjanson. 2009. "Empirical Forecasting of Slow-Onset Disasters for Improved Emergency Response: An Application to Kenya's Arid North." *Food Policy* 34 (4): 329–339.

Naschold, F., and C. B. Barrett. 2011. "Do Short-Term Observed Income Changes Overstate Structural Economic Mobility?" *Oxford Bulletin of Economics and Statistics* 73 (5): 705–717.

OECD (Organisation for Economic Co-operation and Development). 2013. Development Assistance Committee (DAC) Creditor Reporting System Online. Accessed December 19. http://stats.oecd.org/Index.aspx?datasetcode=CRS1.

OFDA (Office of US Foreign Disaster Assistance) and CRED (Centre for Research on the Epidemiology of Disasters). 2013. EM-DAT: The OFDA/CRED International Emergency Disasters Database. Accessed December 19. www.emdat.be/.

Rockmore, M. 2013. "The Cost of Fear: The Welfare Effects of the Risk of Violence in Northern Uganda." Unpublished working paper, Clark University, Worcester, MA, US.

Sen, A. 1981. *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford, UK: Oxford University Press.

Shoham, J., F. Watson, and C. Dolan. 2001. *The Use of Nutritional Indicators in Surveillance*. London: Overseas Development Institute.

Tiwari, S., E. Skoufias, and M. Sherpa. 2013. *Shorter, Cheaper, Quicker, Better: Linking Measures of Household Food Security to Nutritional Outcomes in Bangladesh, Nepal, Pakistan, Uganda, and Tanzania.* Washington, DC: The World Bank.

Torlesse, H., L. Kiess, and M. W. Bloem. 2003. "Association of Household Rice Expenditure with Child Nutritional Status Indicates a Role for Macroeconomic Food Policy in Combating Malnutrition." *The Journal of Nutrition* 133:1320–1325.

USAID (United States Agency for International Development) and Ethiopia, DRMFSS (Disaster Risk Management and Food Security Sector). 2011. *An Atlas of Ethiopian Livelihoods*. Addis Ababa: Livelihoods Integration Unit, DRMFSS, Ministry of Agriculture and Rural Development, and USAID Ethiopia.

Victora, C. G., L. Adair, C. Fall, P. C. Hallal, R. Martorell, L. Richter, and H. S. Sachdev. 2008. "Maternal and Child Undernutrition: Consequences for Adult Health and Human Capital." *The Lancet* 371 (9609): 340–357.

WFP (World Food Programme). 2013. Food Aid Information System (FAIS). Accessed December 19. www.wfp.org/fais/.

WHO (World Health Organization). 2013. Global Database on Child Growth and Malnutrition. Accessed December 19. www.who.int/nutgrowthdb/en/.

Wiesmann, D., L. Bassett, T. Benson, and J. Hoddinott. 2009. *Validation of the World Food Programme's Food Consumption Score and Alternative Indicators of Household Food Security*. Discussion Paper 00870. Washington, DC: International Food Policy Research Institute.

World Bank. 2013. World Development Indicators Online. Accessed December 19. http://data.worldbank.org/data-catalog/world-development-indicators.