The search for the perfect indicator: Reflections on monitoring and evaluation of resilience for improved climate risk management

Stephane Hallegatte, Nathan L. Engle*  

World Bank, 1818 H. Street NW, Washington DC 20433, USA

ARTICLE INFO

Keywords:  
Resilience  
Measurement  
Metrics  
Indicators  
Monitoring & evaluation  
Climate change  
International development

ABSTRACT

Resilience-building activities are increasingly evident in development projects, and as a result, there is a growing focus on monitoring and evaluating the associated outcomes of these projects for improved climate risk management. Significant challenges in measuring resilience, however, contribute to both a tendency towards imperfect quantified metrics, and a quest for universal resilience indicators that can be aggregated across projects, institutions, and geographies. In this paper, we draw from lessons across various sectors typically outside of the traditional resilience and climate risk management realm to highlight potential pitfalls and unintended consequences of such metrics. We then discuss several “thought experiments” to identify the desired characteristics development practitioners would want projects to demonstrate, but for which there exist risks for imperfect aggregated indicators to create perverse incentives. Process-based metrics that focus on the quality of a project’s design and implementation are more likely to avoid these pitfalls and should be considered a viable alternative to aggregated universal resilience indicators.

1. Perspective

There is increasing demand on development practitioners to demonstrate “results” of investments and interventions. This pressure is both well intentioned and necessary, as tax-payer money should be used most efficiently and effectively, and the lessons from successes and failures should be used to improve development aid.

This focus on “results” is also evident in the field of climate risk management, particularly as it relates to the concept of resilience measurement. Here, we use the term “resilience”, as defined by the Intergovernmental Panel on Climate Change (IPCC) in its broadest sense as it relates to managing climate risk to include shocks (such as natural disasters like storms) and stresses (long-term trends such as climate change and desertification). An action to build resilience or contribute to improved climate risk management can do so by reducing exposure (e.g., making sure the population is not affected by building outside of flood zones, reducing the share of climate-sensitive sectors); by reducing vulnerability (e.g., minimizing the damage when the population is affected by a flood with appropriate building norms, or making agriculture climate-smart); or by making the population better able to cope with and recover from shocks and adapt to stresses (e.g., improving access to emergency borrowing or insurance, or improving the ability of workers to shift to sectors that are not affected by climate change) (Hallegatte et al., 2016).

* Corresponding author.

E-mail address: nengle@worldbank.org (N.L. Engle).

1 The International Panel on Climate Change (IPCC) defines resilience as “the capacity of a social, economic, and environmental systems to cope with a hazardous event or disturbance, responding or reorganizing in ways that maintain its essential function, identity, and structure, while also maintaining the capacity for adaptation, learning, and transformation” (IPCC, 2014).
Building resilience is an explicit objective in development strategies and projects, and large amounts have been spent in pursuit of this goal (African Development Bank, 2018). It therefore makes sense to examine closely the resilience-related outcome of these projects. Even when resilience is not a primary objective, projects can support resilience building, and thus should also be evaluated accordingly. Early resilience-related indicators have tended to emphasize “inputs” (OECD-DAC, 1991), for example the finance committed to climate risk management activities. However, there is now a subsequent push to measure the “outcomes,” (i.e. how much climate risk has been reduced or how much resilience has been “produced” with these resources?) (OECD, 2014; Bours et al., 2014b).

These questions are not new to international development. Measuring development effectiveness has been a priority, and a persistent challenge, for decades. However, there are some specific challenges linked to measuring resilience. Resilience is a complex concept with many competing definitions, and its definition is somewhat subjective and dependent on value judgment and preferences. Also, resilience can be directly observed only when a shock occurs, and may depend on the type and magnitude of the shock. As a result, it is difficult to measure and monitor.²

Many feel that we, as development practitioners, will become better able to increase resilience – and through increased resilience, bring improvements to peoples’ well-being – if we can measure it in a quantified way. As the saying goes, “what does not get measured cannot be managed.” In other words, good indicators for the resilience generated by projects should help us learn from successes and failures and prioritize the best resilience-building projects, and in so doing, create appropriate incentives for teams and institutions to ensure that their activities contribute most effectively and efficiently to this objective.

Quantifiable indicators are also sought by the private sector, for example to identify the riskiness of investing in a certain project or set of projects, or defining a preferred asset class that contains more “resilient” investments or more “resilience-building” investments (Koh et al., 2017). There is interest in such metrics from social-impact investors as well, and good resilience metrics would facilitate the mobilization of private capital toward resilience-building projects.

However, the quest for universal resilience indicators has been challenging. Many indicators have been proposed.³ Each has a logical purpose, but in practice, most come with weaknesses; not least of which is the challenge of attributing a given activity or project to resilience-related outcomes that have far longer time horizons than typical development institution project cycles. Still, pressure on development institutions is increasing to design approaches to systematically measure resilience benefits for climate risk management through a single metric or indicator/index that allows for comparability and aggregation of resilience-related outcomes. Particularly worrisome is the fact that any imperfect indicator – with its inability to measure exactly what we want – can easily lead to perverse incentives for practitioners, and favor outcomes that are very different from those intended. It is therefore important to consider the potential pitfalls of over-relying on imperfect quantified indicators to measure progress and to prioritize future investments for resilience building (Leiter and Pringle, 2018).

2. Experience from other sectors

To illustrate, we first draw insights from sectors outside of the traditional resilience and climate risk management realm, where measuring success is also difficult and at least partly subjective, and where quantified indicators have been used for management purposes, but have led to unintended consequences.

2.1. Education

Professionals in the education sector often rely on quantitative metrics, for instance to measure school performance through standardized tests. However, there are well known risks when such metrics occupy too much space in decision-making. For one, a teacher’s role arguably extends beyond what can be easily quantified. Also, attribution is difficult (e.g., a bad score may not be due to a bad teacher or a bad school, but possibly due to disadvantaged backgrounds of the student). The perverse incentive here is that many studies suggest that performance-based pay for teachers has resulted in teachers “teaching to the test,” and ultimately to score inflation (Dee and Jacob, 2011), with little or no demonstrable improvement in education (Menken, 2006; Fuller et al., 2007; Lee and Reeves, 2012; Reback et al., 2014).

2.2. Health care

In health care, the use of quantified indicators has led to similar issues. For instance, hospitals in the United Kingdom once held patients in ambulances outside emergency facilities to reduce their wait time inside the building. They did this to remain below the 4-hour wait time threshold against which hospital performance was measured (Bevan and Hood, 2006). Perhaps even more problematic, poorly designed indicators can push practitioners to avoid the most difficult cases. Since 1992, annual risk-adjusted mortality rates have been public for all hospitals and surgeons providing Coronary Artery Bypass Graft Surgery in Pennsylvania. In a survey conducted in 1996, 59% of the cardiologists reported increased difficulty in finding doctors willing to perform surgery in patients with the highest mortality risk, and 63% of the cardiac surgeons reported that they were less willing to operate on such patients in response to the publication of mortality rates (Schneider and Epstein, 1996).

² Similar challenges are evident with measuring adaptation, as noted in: Bours et al. (2014a), Dinshaw et al. (2014), Leiter (2018), OECD (2008).
³ Examples of such indicators and indices include: BRIC, RIMA, and the Rockefeller Foundation’s City Resilience Index.
2.3. Criminal justice

Similarly, imperfect indicators in the justice system have resulted in undesirable policing behaviors. In France, evidence shows that some police officers shifted their focus from solving crimes to tracking illegal immigrants, as it was an easier way to increase their number of arrests (one indicator considered in performance bonuses). Also in France, the judiciary system has incentivized the reduction of case duration, which has indeed led to a successful decrease in the average length of one’s case. However, a closer look shows that only the longest cases have experienced a reduction in duration, while the duration of short cases has been increased. Overall, most cases have seen an increase in duration, not a decrease (Bacache-Beauvallet, 2011). This shows that indicators can result in a reprioritization of resources, which can represent an inequitable shift without careful attention paid to those that receive fewer resources because of the changes.

2.4. Unemployment services

And finally, when unemployment agencies are evaluated using the share of placed workers, a very reasonable indicator at first sight, there is a tendency for the agencies to shift their efforts on the individuals who are most likely to find a job, at the expense of the most vulnerable people who are arguably in the most need for the services (Anderson et al., 1993).

It is important to note that in the examples above, not all of the indicators were used for direct monetary incentives; some were designed for the sole purpose of monitoring. One common dynamic observed across these cases though is that when an indicator is unable to properly account for the difficulty of an action being incentivized/dis-incentivized (even with risk-based indicators, such in the health care example), professionals tend to focus on what is relatively easier. However, there is often societal benefit in tackling the most difficult cases, for instance, because the worst crimes are often the most difficult to solve. In the case of resilience and climate risk management this extends to wanting to support the people and communities most vulnerable to natural hazards (i.e., those living in low-capacity environments where projects are particularly challenging to implement, and where data and capacity to track progress are often harder to come by).

Another concern is how complementing intrinsic motivation (i.e., the willingness of individuals to do their job properly) with extrinsic motivation (e.g., a monetary reward based on performance) can counterintuitively reduce the intrinsic motivation, to the point where the overall motivation is reduced by the additional monetary incentive (Bénabou and Tirole, 2006; Falk and Kosfeld, 2006). For example, one study shows that when day care providers introduced a fee for parents when they arrive late to pick up their children, parents tended to arrive later more often (Gneezy and Rustichini, 2004). Because they pay when they are late, the moral imperative of being on time becomes less important to the parent, and the net impact is to reduce the incentive to be on time. The same effect has been observed for the willingness of local communities to accept undesired local projects, such as airports, new chemical or nuclear plants, or prisons (Frey and Oberholzer-Gee, 1997). Providing monetary compensation to the people living close these projects can reduce their willingness to accept such a project for moral and ethical reasons (e.g., contribution to society or their community), and can make people even more reluctant to accept the project than without compensation.

As we develop indicators to measure the resilience benefits of our projects, the lessons from these sectors are important to keep in mind: we should not forget that indicators create incentives, and that bad indicators create bad incentives (Heckman et al., 2011). While an imperfect indicator could appear “good enough” to track the performance of a project or an institution, it can be dangerous if staff start to focus their efforts on “looking good” according to the indicator. Like the teachers who “teach to the test,” one could easily imagine the scenario where development practitioners start to “develop to the indicator”, with potentially negative impacts on peoples’ lives, as well as and reduced efficiency in the use of scarce resources. This effect has been referred to as the Campbell’s law, which states that “the more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor” (Campbell, 1979).

3. How to ensure that indicators create the right incentive?

Now, we turn to a few “thought experiments” that further illustrate the risk for an indicator to create perverse incentives for development professionals and institutions operating in the area of climate risk management. In doing so, we hope to raise the awareness as to how an indicator might realistically affect incentives and decision-making, to ensure that it does not lead us toward the wrong projects.

Each experiment considers one or several alternatives for a project, with one clearly superior to the other(s), and explores how different indicators can create an incentive for selecting a suboptimal solution. The experiments cover seven characteristics (see Fig. 1) that we purport would be desirable for projects to demonstrate (note this is not a complete nor consensus list of characteristics): (1) efficient; (2) context specific; (3) fair; (4) transformational; (5) comprehensive; (6) robust; and (7) difficult.

(1) Efficient. Consider a new road that will be built in a low-income country. A climate risk screening exercise shows that the road will be exposed to significant flooding. Two options are on the table to manage this climate risk. Option A reinforces the road and increases drainage capacity, which makes the road more robust but increases its cost by 30%. Option B moves the road by 2 km so that it avoids the flood zone, at a marginal additional cost. Here, the question is whether the indicator will favor the costly solution at the expense of the more efficient one. Input-based indicators (e.g., “how much of the project finance can one attribute to climate change adaptation or resilience?”) risk favoring and incentivizing the more expensive options.

(2) Context specific. Consider two communities in very different situations. Community A suffers from minor droughts because it has...
minimal water storage potential and is completely reliant on rainfall for farming. Farmers in this community cultivate crops that can survive drought, but have low productivity. As a result, poverty is widespread in community A. Community B is richer and uses irrigation massively, applying groundwater pumping to provide the water, which leads to rapid salinization of the water reserves and creates a threat to longer-term water supply for agriculture and human consumption. Here, the question is whether universal resilience indicators can capture the nuances that community A can become more resilient due to investments in irrigation and move toward more water-intensive crops, while community B can become more resilient due to dis-investment in irrigation and a move to less water-intensive agriculture. Similar issues will arise with adaptive social safety nets, which may either make a community sustainable and resilient by facilitating coping during a bad year, or in other situations may keep people in locations where they have no credible long-term prospects, pathways to prosperity, and decent living conditions (Hallegatte et al., 2015). Positive lists of adaptation action (e.g., all irrigation projects count as adaptation to climate change) will struggle to account for the context of the operation and may lead to favoring one-size-fits-all solutions, and in some cases, maladaptation (when the resilience-building activities are considered across broader spatial scales and longer temporal dimensions).

(3) Fair. Imagine two coastal protection projects, each project costing $100 million. Project A is expected to prevent $20 million in annual flood losses, while project B is expected to prevent $10 million in annual losses. On the basis of a classical cost-benefit analysis, project A appears superior. However, the difference in benefits arise from the fact that project A protects a very wealthy neighborhood, while project B helps an informal settlement where people are poor and few tangible assets to protect. While the rich neighborhood could finance its protection itself (but is happy not to if somebody else pays for it), the informal settlement cannot self-finance its protection. Furthermore, while project A reduces asset losses from floods without many development benefits, project B helps many people escape poverty over time, as they can spend less of their little savings on repairing their homes on a regular basis. Thus, indicators for resilience based on “avoided losses” or monetary benefits anticipated by the project might favor projects similar to project A (and more generally any projects protecting richer regions and individuals, which concentrate most of the economic value in a country). Resilience indicators thus need to be subtle enough to include equity and poverty considerations, and should look beyond direct benefits to consider the development dividends from different operations (Hallegatte et al., 2016). Otherwise, projects will be incentivized that might not produce the best dividends for the intended beneficiaries.

(4) Transformational. Think of a coastal protection investment project at an early stage of project design. It targets a city where roughly 100,000 people are flooded every year or every other year. Considering available resources, two options are on the table: focusing investments in the highest risk areas, which cover 10,000 people, and providing them with protection against the 100-year event (i.e., the event with a 1% probability of occurrence every year); or spreading investments over the whole population, which with the same resources marginally reduces flood frequency, from one in two years to one in five years, but for the whole population of 100,000. Indicators that focus on the number of beneficiaries (e.g., number of people with increased resilience) could very well drive business toward projects with marginal impacts but many beneficiaries. As a result, it becomes less likely that these projects will be transformational and help people escape poverty for good. Instead, indicators such as these should take into account not only the number of beneficiaries, but also the extent to which the project can change lives and generate sustained benefits.

(5) Comprehensive. Imagine an urban development project in a large coastal city. A risk screening process flags, unsurprisingly, that climate change and disaster risks need to be considered in the design of the project. In response, the team adds a stand-alone component to the project that finances dikes and other coastal protection infrastructure. Here, the danger is that adding a separate component may give the impression that the problem is solved and the climate risks will be effectively managed. And indeed, this “simple action bias” is a well-documented behavioral bias (Weber and Johnson, 2011; Weber et al., 2008). That is, when confronted with a risk or a problem, we tend to do one minor thing and consider the problem solved. Similarly, a team working on an urban project may decide to add a couple of sea walls, and will assume that all flood related risks have been addressed. But, disaster and climate risks are sometimes so significant that they cannot be simply managed by a separate component; they need to be accounted for in the design across the full project. Thus, teams should ensure not to push the clients toward singular actions or box-ticking, rather incentivizing more significant resilience-building interventions. This logic also cascades up to the need, oftentimes, to account for the risks beyond the project level. That is, the risks and resilience-building

---

**Fig. 1.** Examples of desirable characteristics for projects that are potentially compromised without careful consideration of the pitfalls and limitations of resilience indicators.
measures developed on a project-by-project basis might not be adequately accounting for the relationships with other activities occurring within the country, nor the implications of the resilience-building measures beyond the individual project cycle. Instead, desirable resilience indicators would need to be sufficiently situated in country planning processes and broader programmatic approaches for investments, and would need to manage the multiplicity of scales, from the national to the project-level. Thus, a single metric or aggregated outcome indicator/index is likely to fail in this respect, with process-oriented metrics that characterize the quality of climate risk screening and management being more suited for such a task.

(6) Robust. Consider a hydropower investment that is highly dependent on water availability and rainfall (Conway, 2017). The team hires a climate science consultant to provide information about future climate conditions at the location of the dam to inform the project design. The consultant uses one of the leading climate models, which projects an increase by 35% of rainfall at the chosen location, and the project design is adjusted accordingly. Since climate change has been taken into account in the project design, many indicators would count the hydropower dam as “resilient” or “climate informed.” But then of course, the consultant may have used a different leading climate model, and the projected change in rainfall could have been a 20% decrease, which would have led to a completely different design. Future climate conditions are highly uncertain, and there are regions, like West Africa or India, where models even disagree on whether rainfall will increase or decrease in the future (Cervigni, 2015). Indicators that label a project as resilient that has been designed using climate model outputs, without careful consideration, may encourage teams to design their project using a preferred climate model that labels it “climate informed” at the lowest possible cost, without considering the uncertainty in future climate conditions. Resilience indicators should incentivize projects that are “robust” (i.e., projects that deliver development benefits under a wide range of possible climate and socio-economic conditions).

(7) Difficult (enough). Finally, take two projects in a country, both aiming to improve the resilience of agriculture production to drought in different communities. The first project takes place in a community with high capacity, and its chances of success are close to 100%. The second project targets a very poor and vulnerable community with low capacity, which would have a major impact on their livelihoods. The latter project, however, faces political and technical obstacles and its chances of success are estimated at only 25%. Indicators that do not account for the level of risk of the intervention could easily create an incentive to focus on safe (but marginal and less impactful) projects, at the expense of more ambitious and transformational ones. However, if we try to correct for the bias towards easier projects by explicitly favoring risky projects, development practitioners would fall into the opposite problem of favoring projects with excessive risk levels. Thus, development institutions must strike a careful balance between supporting the difficult and more complex projects with scarce concessional finance and avoiding the acceptance of excessive risks that could plague the project.

This problem is magnified by the challenge of attributing impacts to a project, especially in fast-changing areas of the world and in the presence of many parallel policy changes and public and private actions. Consider one resilience-building project – say the retrofitting of all hospitals to enable them to resist storms and earthquakes – that can be implemented in one of two regions in a country. Region A has seen a lot of progress in terms of resilience in the last decades, supported by targeted and efficient action by the government. Region B is fragile, with recent conflicts, and has seen limited action from local authorities to support climate risk management. If the indicator can measure resilience perfectly but cannot attribute resilience gain to a specific intervention, it would incentivize actions in region A at the expense of the arguably more-in-need region B. Indeed, 5 years after the project is completed, it is likely that resilience in region A will have kept improving, making it possible to make a loose connection with the project and the observed project. This might also incentivize teams to support projects that would have been implemented anyway (i.e., projects that crowd out domestic actions and are not “additional” compared to a no-intervention scenario). To help avoid this bias, projects should articulate a “theory of change” (i.e., the logic model that illustrates the causal link between the intervention and the gains in climate risk management, or the project’s contribution to a resilience pathway), and project-specific indicators able to measure the existence and magnitude of this link along the results chain (World Bank, 2017).

These thought experiments do not comprehensively illustrate all of the challenges and drawbacks that universal resilience indicators may create. But being aware of these risks can help development practitioners working in the realm of climate risk management avoid some pitfalls.

What can be the way forward? Some of these risks and problems can be mitigated by improving the indicator(s), and ensuring that they can capture the various dimensions discussed here. For instance, an indicator measuring the number of resilience beneficiaries could also include thresholds for how much must be gained in that indicator for a beneficiary to be counted. To ensure projects targeting the poorest and most vulnerable people are not at a disadvantage – even though the monetary benefits from these projects may be smaller and their implementation risks larger – individual benefits accruing to the poor and rich can be valued differently, as the World Bank did in the Unbreakable report, to measure socioeconomic resilience.

Risks created by imperfect indicators can also be reduced by combining several indicators. If an indicator is based on aggregated benefits, and therefore risks favoring better-off beneficiaries over poorer ones as highlighted in the example above, then a complementary indicator could be the same measure of aggregated benefits, but only counting people that are in the bottom 20% in terms of income. While introducing multiple indicators may appear complex (and is sometimes costlier in terms of time and resources for operational teams), a set of complementary indicators is less likely to lead to large negative outcomes than a single indicator.

Even more appropriate would be to develop process-based metrics to measure the extent to which projects are: (1) designed to account for climate risks and other uncertainties and are flexible enough in their implementation and monitoring and evaluation (M&E) system to be able to adjust as climate risks ensue; and (2) support resilience building outcomes in a given community, ecosystem, or country. Both could require that a range of climate models are used to stress test the project and identify the most appropriate activity or investment – based on strict procedural norms regarding how these different climate model outputs are taken into account.
to reach a robust decision. This would certainly benefit from clear and detailed guidance on the criteria for designing such projects in given sectors.

Regardless of the quality of an individual indicator and any associated complementary indicators, it is evident that aggregated quantitative resilience metrics will only take us so far. Resilience is as much about infrastructure and financial instruments as it is about governance, voice, and empowerment. But governance, voice, and empowerment are not easy to quantify and measure, and should not be ignored at the expense of the search for quantified metrics.

In the face of the complexity of the issue, making development more resilience-oriented will require that the tools used for project prioritization, design, and M&E have enough flexibility to include resilience in the most relevant ways. This will likely come at the expense of aggregated resilience outcome metrics intended to measure the resilience generated by a portfolio of projects that are different in nature. Given the dangers identified here, we feel that this is a worthwhile tradeoff for the development community to make.

Conflict of interest

The authors declare that there is no conflict of interest.

References

African Development Bank (AfDB), Asian Development Bank (ADB), European Bank for Reconstruction and Development (EBRD), European Investment Bank (EIB), Inter-American Development Bank Group (IDB/G), Islamic Development Bank (ISDB), World Bank Group (WBG), 2018. Joint Report on Multilateral Development Banks’ Climate Finance, 2017. http://www.ebrd.com/2017/joint-report-on-mdbs-climate-finance.

Anderson, Kathryn, Burkhauser, Richard, Raymond, J., 1993. The effect of creaming on placement rates under the job training partnership act. Ind. Labor Relat. Rev. 46 (4), 613–624.

Bacache-Beauvallet, M., 2011. Rémunération à la performance. Actes de la recherche en sciences sociales 4, 58–71.

Bénabou, R., Tirole, J., 2006. Incentives and prosocial behavior. Am. Econ. Rev. 96 (5), 1652–1678.

Bevan, G., Hood, C., 2006. What’s measured is what matters: targets and gaming in the English public health care system. Public Admin. 84 (3), 517–538.

Bours, D., McGinn, C., Pringle, P., 2014. Guidance Note 1: Twelve Reasons Why Climate Change Adaptation M&E is Challenging. SEA Change Community of Practice, and UK Climate Impacts Programme.

Bours, D., McGinn, C., Pringle, P., 2014b. Selecting indicators for climate change adaptation programming: guidance for M&E of climate change interventions series. Guidance Note 2.

Baseline Resilience Indicators for Communities (BRIC). http://artsandsciences.sc.edu/geog/hvri/bric.

Campbell, D.T., 1979. Assessing the impact of planned social change. Eval. Program Plann. 2 (1), 67–90.

Cervigni, R. (Ed.), 2015. Enhancing the Climate Resilience of Africa’s Infrastructure: The Power and Water Sectors. The World Bank.

Conway, et al., 2017. Hydropower plans in eastern and southern Africa increase risk of concurrent climate-related electricity supply disruption. Nat. Energy 946–953.

Dee, T.S., Jacob, B., 2011. The impact of No Child Left Behind on student achievement. J. Policy Anal. Manage. 30 (3), 418–446.

Dinsah, A., 2014. Monitoring and Evaluation of Climate Change Adaptation: Methodological Approaches, OECD Environment Working Papers, No. 74. OECD Publishing.

Falk, A., Kosfeld, M., 2006. The hidden costs of control. Am. Econ. Rev. 96 (5), 1611–1630.

Frey, B.S., Oberholzer-Gee, F., 1997. The cost of price incentives: an empirical analysis of motivation crowding-out. Am. Econ. Rev. 87 (4), 746–755.

Fuller, B., Wright, J., Gesicki, K., Kang, E., 2007. Gauging growth: How To Judge No Child Left Behind? Educ. Res. 36 (5), 268–278.

Gneezy, U., Rustichini, A., 2004. Incentives, punishment and behavior. Adv. Behav. Econ. 572–589.

Hallegatte, S., Bangalore, M., Bonzani, L., Fay, M., Kane, T., Narloch, U., Vogt-Schilib, A., 2015. Shock Waves: Managing the Impacts of Climate Change on Poverty. The World Bank.

Hallegatte, S., Vogt-Schilib, A., Bangalore, M., Rozenberg, J., 2016. Unbreakable: Building the Resilience of the Poor in the Face of Natural Disasters. World Bank Publications.

Heckman, J.J., Heinrich, C.J., Smith, J., 2011. Performance standards and the potential to improve government performance. In: Heckman, J.J., Heinrich, C.J., Courty, P., Marschke, G., Smith, J. (Eds.), The performance of performance standards. W.E. Upjohn Institute for Employment Research, 1-14, Kalamazoo, MI.

IPCC, 2014. Climate change 2014: synthesis report. In: Core Writing Team, Pachauri, R.K., Meyer, L.A. (Eds.), Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland, pp. 15.

Koh, J., Mazzucato, E., Trabacchi, C., 2017. An Investor Guide to Physical Climate Risk & Resilience: An Introduction. Global Adaptation & Resilience Investment

Leiter, T., Pringle, P., 2018. Pitfalls and potential of measuring climate change adaptation through adaptation metrics. In: Christiansen, L., Martinez, G., Naswa, P. (Eds.), Adaptation Metrics: Perspectives on Comparing, Measuring and Aggregating Adaptation Results. UNEP DTU Partnership, Copenhagen, Denmark.

Leiter, T., Pringle, P., 2018. Pitfalls and potential of measuring climate change adaptation through adaptation metrics. In: Christiansen, L., Martinez, G., Naswa, P. (Eds.), Adaptation Metrics: Perspectives on Comparing, Measuring and Aggregating Adaptation Results. UNEP DTU Partnership, Copenhagen, Denmark.

Menken, K., 2006. Teaching to the test: How No Child Left Behind student achievement. J. Eval. Policy Anal. 34, 209–231.

Reback, R., Rockoff, J., Schwartz, H., 2014. Under pressure: Job security, resource allocation and productivity in schools under No Child Left Behind. Am. Econ. J.: Econ. Policy 6, 207–241.

Food and Agriculture Organization of the United Nations’ Resilience Index Measurement and Assessment (RIMA) Model http://www.fao.org/3/a-i1402e.pdf.

Rockefeller Foundation’s City Resilience Index https://www.rockefellerfoundation.org/report/city-resilience-index/.

Schneider, E.C., Epstein, A.M., 1996. Influence of cardiac-surgery performance reports on referral practices and access to care—a survey of cardiovascular specialists. N. Engl. J. Med. 335 (4), 251–256.

Weber, E.U., Johnson, E.J., 2011. Psychology and Behavioral Economics Lessons for the Design of a Green Growth Strategy: White Paper for World Bank Green Growth Knowledge Platform.

Weber, E.U., Lindemann, P.G., Plessner, H., Betsch, C., Betsch, T., 2008. From Intuition to Analysis: Making Decisions with Your Head, Your Heart, or by the Book. World Bank, 2017. Operational guidance for monitoring and evaluation (M&E) in climate and disaster resilience-building operations. World Bank Group. http://documents.worldbank.org/curated/en/692091513937457908/pdf/122226-ReME-Operational-Guidance-Note-External-FINAL.pdf.