

INDEPENDENT REPORTING MECHANISM

BEYOND THE BASICS OGP ACTION PLANS 2012–2015



Second IRM Technical Paper
Fall 2016
Joseph Foti, IRM Program Director



INDEPENDENT REPORTING MECHANISM BEYOND THE BASICS OGP ACTION PLANS 2012-2015



| | |
|---|----|
| SUMMARY OF MAJOR FINDINGS | 3 |
| 1 INTRODUCTION | 5 |
| WHAT WOULD PROGRESS LOOK LIKE? | 5 |
| ABOUT THE DATA..... | 6 |
| BOX: WHAT IRM REPORTS CAN AND CANNOT ASSESS ABOUT THE OGP THEORY OF CHANGE | 7 |
| II OGP PROCESSES | 9 |
| OBSERVATIONS..... | 10 |
| PERFORMANCE IMPROVEMENTS | 10 |
| BEYOND COMPLIANCE: TOWARD MEANINGFUL PARTICIPATION..... | 12 |
| III ACTION PLAN CONTENT AND IMPLEMENTATION..... | 15 |
| COUNTING STARS: FROM ACTION PLAN “LEANNESS” TO “MODEL” COMMITMENTS | 15 |
| WHY AREN’T THERE MORE STARS? | 17 |
| WHAT CAN STARRED COMMITMENTS TELL US ABOUT THE DIRECTION OF CHANGE?..... | 17 |
| DIVERGENCE IN ACTION PLAN “LEANNESS” | 18 |
| FALLING STARS: UNDERSTANDING CHANGES ACROSS ACTION PLANS..... | 20 |
| MOVING TOWARD ACCOUNTABILITY?..... | 22 |
| IV TYING IT TOGETHER: ACTION PLANS AND NATIONAL- LEVEL CHANGE | 25 |
| THE RELATIONSHIP BETWEEN BETTER PARTICIPATION AND IMPLEMENTATION..... | 25 |
| BEYOND THE ACTION PLAN: NATIONAL-LEVEL CHANGE | 26 |
| V OPERATIONAL RECOMMENDATIONS..... | 31 |



SUMMARY OF MAJOR FINDINGS

OPEN GOVERNMENT PARTNERSHIP

BEYOND THE BASICS

OGP ACTION PLANS 2012-2015

The Independent Reporting Mechanism (IRM) Second Technical Paper reports on major trends and changes in Open Government Partnership (OGP) action plans. This paper compares, for the first time, how governments have performed from their first to second action plans.

OGP action plans can be evaluated on the following criteria:

- **Process:** How open, participatory, and meaningful is government-civil society dialogue in developing action plans?
- **Commitments:** How clear, relevant, and ambitious were action plans? Were ambitious commitments implemented? If not, why?
- **Impacts:** Can we see linkages between OGP action plans and major changes in governance?

Process: Action plan creation has improved overall, but many countries lag and consultation is often shallow

Most countries have improved in the formal requirements of consultation from action plan to action plan. In fact, nearly all countries now have in-person consultations in the development of the action plan.

This promising finding does not, however, speak to the quality of dialogue or the depth of co-creation of action plans. When we look at the degree of influence the public might have on the contents of an action plan, the picture is more stark. Most countries have opportunities for public input, but less than half have opportunities for agenda setting and iterative dialogue. This is a major area for continued improvement.

Commitments: Action plans are more measurable, relevant, and implemented than ever before, but ambition is still lacking

Less than five percent of all commitments were “starred.” That is, these commitments were (1) specific and measurable, (2) clearly relevant to OGP, (3) marked as having “transformative” potential impact, and (4) saw “significant” or better progress toward completion. The biggest barrier to becoming a starred commitment is inadequate ambition, followed closely by weak rates of completion.

Over time, the average commitment (and action plan) has improved in specificity, relevance, and completion. However, potential impact (a component of ambition) declined. This can in part be attributed to an adjustment of the IRM method, but it may also be a sign of governments scaling back ambition in an attempt to complete commitments. It is important to note that the OGP Support Unit and multilateral partners have put concerted efforts into improving specificity and relevance. Further investment will need to be made to ensure that action plans demonstrate adequate ambition and that rates of completion—especially of ambitious commitments—rise.

An additional finding shows that there is a growing divergence between high-performing OGP countries and more poorly performing OGP countries. Specifically, at the top a growing group of countries has increased the percentage and number of strong commitments, while at the bottom another growing group still struggles with the basics of commitment design and implementation. This results in a shrinking middle of countries.

Impacts: While macro-level changes have occurred, the biggest achievements have occurred where there is Support Unit technical support and where government staffing remains stable

While there is wide variation among OGP countries performance on key OGP indicators, it is too early to identify any effects of OGP on third-party indicators, such as budget transparency, civil liberties, or improved access to information. This may be due to the inevitable lag time between action plans and national changes, the possible misalignment of action plans with these major indicators, and the general “stickiness” of indicators like civil liberties scores (such as Freedom House or Economist Intelligence Unit [EIU] civil liberties scores).

Additionally, there is little compelling evidence to suggest that political transitions or particular institutional arrangements contribute to higher rates of completion or more stars. Nonetheless, there is statistically significant evidence to suggest that stability of the civil service office in charge of OGP is positively correlated with greater rates of implementation and higher percentages of starred commitments.

I | INTRODUCTION

After five years of the Open Government Partnership, it is time to ask how far the organization has come and where it should go next.

In 2014 when the first major round of IRM reports was released, we, the IRM staff, wrote the first IRM Technical Paper, which looked at what it meant for a country to be “successful” in OGP. Regardless of definition, success was not as widespread as we had hoped. We found consistent weaknesses in process, quality of action plans, and low ambition of commitments. This was consistent with expectation: when the partnership began, there was little well-publicized guidance on how to develop an action plan, a lack of clarity on what “open government” meant, and few structures established at the national level to coordinate and motivate major open governance reforms. Nonetheless, the technical paper established a valuable baseline for performance across key indicators for OGP action plan development and implementation. OGP’s various stakeholders can now see how we have, and have not, progressed from that initial baseline, and whether the questions asked in the last paper are still the correct questions to ask.

WHAT WOULD PROGRESS LOOK LIKE?

The simple definition of progress that we will use is: “More and more OGP processes and OGP commitments are successful.” So, what is success? The last IRM Technical Paper defined “success” in OGP as follows (from most instrumental to most substantive):

- **Participatory process**—Each action plan would be developed, implemented, and assessed through a participatory process that follows the guidelines set out by the OGP Articles of Governance and the OGP Steering Committee and involves civil society.
- **High-quality commitments implemented**—Commitments would be specific, measurable, and ambitious for the national context in which they are implemented.
- **High-impact**—National participation in OGP (including but not limited to the IRM) would result in more open government. This, in turn, would change lives.

The IRM data can evaluate success across the first of these two elements, as they are (a) directly attributable to OGP, (b) operationalizable, (c) measurable (through reasonable proxies), and (d) available. We are limited in what we can say (in this paper, at least) about the third element, impact. The IRM is not an impact assessment process. Consequently, this paper and the data do not make any claims to the ultimate impact of a context-specific, multifaceted process like OGP.

If we define success in OGP on these aspects—process, commitments, and impacts—then it follows that progress equates to improvement on key indicators of each of these. Improvements in process, in the quality of inputs (commitments), and outcomes (implementation) can be compared across three dimensions, with each answering a separate question.

1. **Improvement across time**—Is OGP improving overall across time?
2. **Across subsequent action plans**—Do individual countries learn to do OGP better?
3. **For all new action plans**—When a new country joins OGP now, are they more successful than when a country joined before?

This paper makes a first attempt to present what the data show us about each of these elements, to offer up some tentative explanations, and, as appropriate, to make a number of operational recommendations for the IRM, the OGP Support Unit, and the Steering Committee.

ABOUT THE DATA

This paper focuses on what data generated through the IRM reporting process can tell us about the health of OGP and is derived from a database of nearly 2,000 assessed commitments. There are three limitations related to the IRM data that are important to understand: sources and method, units of analysis, and timing.

First, a word on sources. The IRM is an accountability mechanism that first and only produces data as a byproduct—a positive externality for its twin goals of creating public accountability and learning to improve the action planning process.

Further, the research methods of the IRM do change over time. This reflects learning by the IRM on how to better carry out its mandate to hold governments accountable. Early versions of the method did not adequately capture the ambition of commitments or civil society engagement. Again, the IRM is an accountability mechanism that creates data useful for social science and is not a social scientific instrument. Nonetheless, the information goes through multiple stages of peer review, and OGP's IRM staff work to create consistency and comparability between action plans and countries. If the changes in research methods might affect the production or interpretation of data, we will make every effort to note that in the text of this paper.

Second, a word on units of analysis. Each IRM report is focused on the process, institutional arrangements, and commitments that make up an action plan. Because of this, readers should not expect the IRM data to reflect the totality of regress and progress on open government, however defined, in a given country. The bulk of each national IRM report focuses on how a country is doing in OGP, not how it is doing in open government.

The data in this paper are based on the two IRM databases, and the two databases are distinguished based on units of analysis. One on process and institutions is measured at the country level. The second focuses on individual commitment-level data. Readers would take care to note that any generalizations about “percentage” of commitments or comparisons should be seen as shorthand—each commitment is not necessarily comparable with other commitments (in any number of attributes such as weight, cost, word count). Nonetheless, we assume the reader will agree that it is the best available data and that the downside risks of not using the data outweigh those of using it.

Third, note that there is an inevitable issue of timing with the IRM process. First, the IRM reports are published seven months after the first year of the action plans. There are problems with the time lag in doing analysis. For instance, the most recent event described by this paper took place in June of 2015, and the most recent action plan development process described took place in 2014, only two years after most countries joined OGP. Consequently, it does not cover the whole of the narrative arc of OGP. We will have to wait another two years for that full picture.

Further, the duration of early action plans and commitments varied. During the first round of OGP action plans, there was inconsistent (or sometimes no) public guidance or “rules of the game” on the process, calendar, and nature of commitments. Some action plans lasted one year (Moldova, El Salvador, and Indonesia), while others last three or more years (Canada and Azerbaijan). Indeed, some action plans had no discernable time frame at all. This is not a cause for calamity; it is normal for any new partnership to move from a period of initial enthusiasm to a period of clarification and consolidation. It does, however, limit the degree to which we can make definitive conclusions and comparisons across time, especially when discussing any aspect concerning completion of commitments. For those action plans evaluated prior to 2015, some (those of one year or less) were evaluated post hoc, while others were evaluated mid-term. For those from 2015 onward, those action plans were evaluated mid-term as all action plans are now two-year action plans. (Note: commitments vary in length from one day to well beyond two years.) Beginning in 2016, the IRM now carries out both mid-term and end-of-term evaluations of commitments. In early 2017, we will publish a synthesis of end-of-term evaluations. Until then, despite inconsistencies in the time frame, we hope the reader will agree that carrying out the analysis on the data, however flawed, is better than not carrying out the analysis. We will make every effort to note when this issue could affect the data.

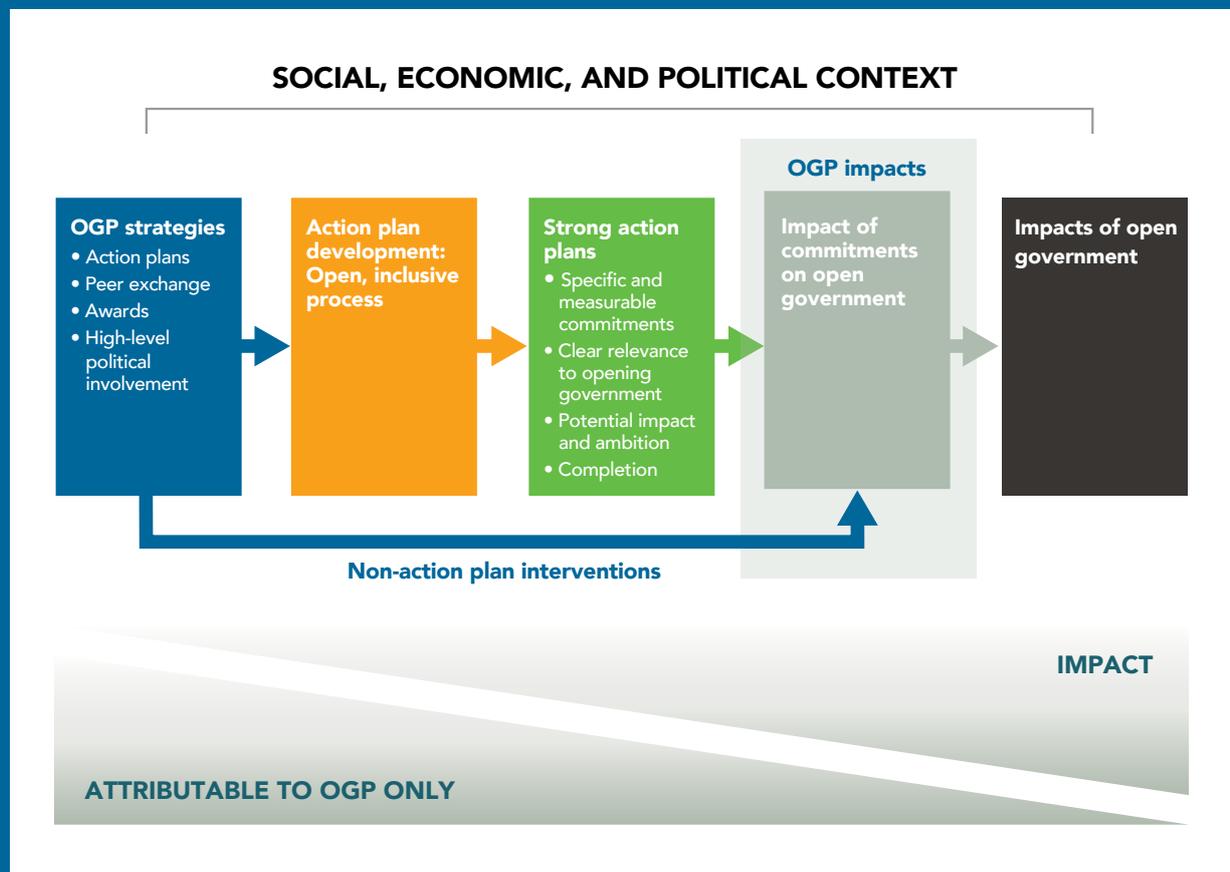
The box, “What IRM reports can and cannot assess,” summarizes the methodological limitations of the IRM with regard to OGP effectiveness.

What IRM Reports Can and Cannot Assess About the OGP Theory of Change

Although they contain a limited amount of contextual information, IRM reports only address OGP action plan development and implementation. Of course, OGP is more complicated than that, and there are broader strategies that go beyond action plans. In attempting to change the relationship of citizens to their governments, OGP employs other methods, such as holding regular global and regional meetings, initiating peer exchanges, garnering multilateral support, and engaging high-level political support. These latter strategies are often reflected in OGP action plans indirectly, but they presumably result in changes to government practice and broader changes in political discourse. Additionally, external factors like civil society capacity, election cycles, and government capacity contribute to progress (or regress) on open government and its ultimate impacts.

The figure below shows what the IRM reports can assess with regard to OGP two-year action plans, but any full assessment of OGP and its effectiveness should keep in mind the broader strategies of OGP as well external factors. When such factors are taken into account, direct attribution becomes more tenuous. A dose of humility is in order in making claims like “This reform would not have happened but for OGP,” or the counterfactual, “This reform would have happened anyway.” Instead, more realistically, one might assess whether OGP action plans (1) contributed to or accelerated implementation or (2) shifted discourse.

Figure: The relationship between attribution and impact in OGP



II | OGP PROCESSES

Countries have mostly improved in complying with basic OGP process requirement, but significant room for improvement remains in improving quality and openness of dialogue.

OGP action plans are to be developed in a participatory fashion, according to OGP's Articles of Governance. There are normative and instrumental arguments for this requirement. From a normative perspective, OGP is supposed to be a model of the type of participatory decision making that it promotes. The OGP theory of change, as articulated in the Four-Year Strategy, reflects a more instrumental approach. The strategy states that meaningful participation strengthens the chance of getting ambitious reforms into action plans. The IRM does not weigh the validity of either of these arguments but, instead, monitors participation during development and implementation of the action plan.

There is strong evidence to show that OGP processes have increasingly conformed with the *letter* of OGP requirements, although IRM data cannot speak to whether participation was carried out in the *spirit* that those requirements were written in. In fact, improvement on formal adherence to the requirements for participation is the clearest indicator of progress among the various areas the IRM can measure. Because of the difficulty in assessing whether consultation was carried out in the spirit of collaboration, we introduce here new indicators to better describe the quality and openness of the consultation.

The OGP Articles of Governance (Addendum C) lay out a number of required steps for consultation. (Note that these are under consideration for revision now.) They are as follows:

- **Availability of timeline:** Countries are to make the details of their public consultation process and timeline available (online at a minimum) prior to the consultation;
- **Adequate notice:** Countries are to consult the population with sufficient forewarning;
- **Awareness raising:** Countries are to undertake OGP awareness-raising activities to enhance public participation in the consultation;
- **Multiple channels:** Countries are to consult through a variety of mechanisms—including online and in-person meetings—to ensure the accessibility of opportunities for citizens to engage;
- **Breadth of consultation:** Countries are to consult widely with the national community, including civil society and the private sector, and to seek out a diverse range of views; and
- **Documentation and feedback:** Countries are to make available online a summary of the public consultation and all individual written comment submissions.

The IRM tracks each of these indicators for all action plans submitted after November 2011, when the Steering Committee adopted these standards.

The prior technical paper showed very little conformity to the requirements during the first round of action plans. This was likely due to the rushed nature of the first action plans, little publicity of the rules (the Support Unit was fewer than five staff members at the time), and lack of clarity around requirements. Given the dismal early performance on these criteria, we ask three questions about the requirements for participation in OGP.

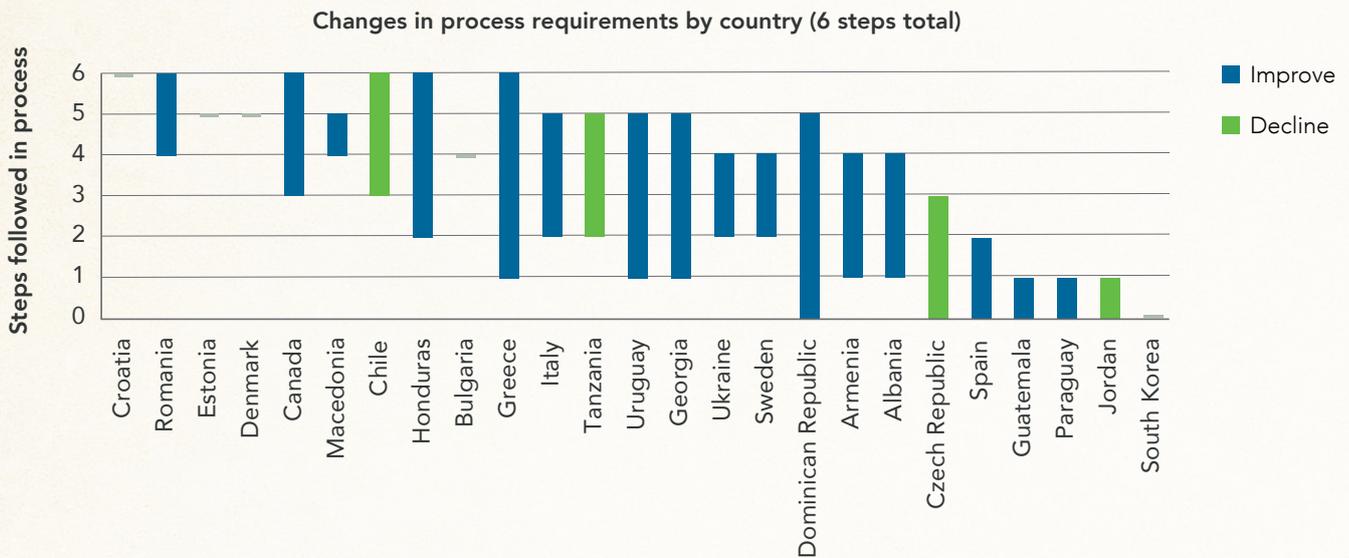
- Did countries improve from action plan to action plan?
- Do new countries perform better the first time around now?
- Has conformity with requirements gone up over time?

OBSERVATIONS

Performance improvements

There is evidence that governments are becoming better at conforming with the established standards. In fact, 26 of 30 countries that have been evaluated twice did the same or better in the second action plan. Only four countries declined. Figure 1 shows this change.

Figure 1: Governments generally improved on process requirements between action plans



In this area, the direction of change is clear. Countries have learned to follow the steps outlined in the Articles of Governance more closely. Nonetheless, two caveats are important. First, most readers would not weigh all six of the steps equally. Some indicators, such as the existence of in-person consultation, are arguably more important than the advance publication of a timeline. Second, these indicators cannot show how meaningful this participation may be. We try to address each of these issues later in the paper by looking at other key information.

Significant time and energy has been given to bringing new countries into OGP. Are new countries doing better? If we limit our evaluation only to new countries, the answer is a clear yes. Figure 2 below gives rates of conformity with the required steps by year. As time progresses, action plan development follows more and more steps. (While the direction of change is clear overall, the year can only explain 5% of variance at the country level: $r^2=.053$. There remains more variation within each cohort than between cohorts.)

Figure 2. Improvement in procedural conformity by year

| | PERCENT OF POSSIBLE STEPS BY YEAR | | |
|------------------|-----------------------------------|------|------|
| Year | 2012 | 2013 | 2014 |
| % possible steps | 43% | 52% | 60% |

Figure 2 above does not show what parts of the process see the highest rates of conformity to guidance. That information is illustrated in Figure 2 below. All but one of the steps saw improvements, with “in-person dialogue” having the highest showing overall. The exception was the declining number of countries with “ongoing forums.” Again, this could be because the Support Unit did not make guidance widely available until 2015 (and the period assessed was 2014).

Figure 3. Improvements and declines in process requirements between NAPs 1 and 2 (n=25)

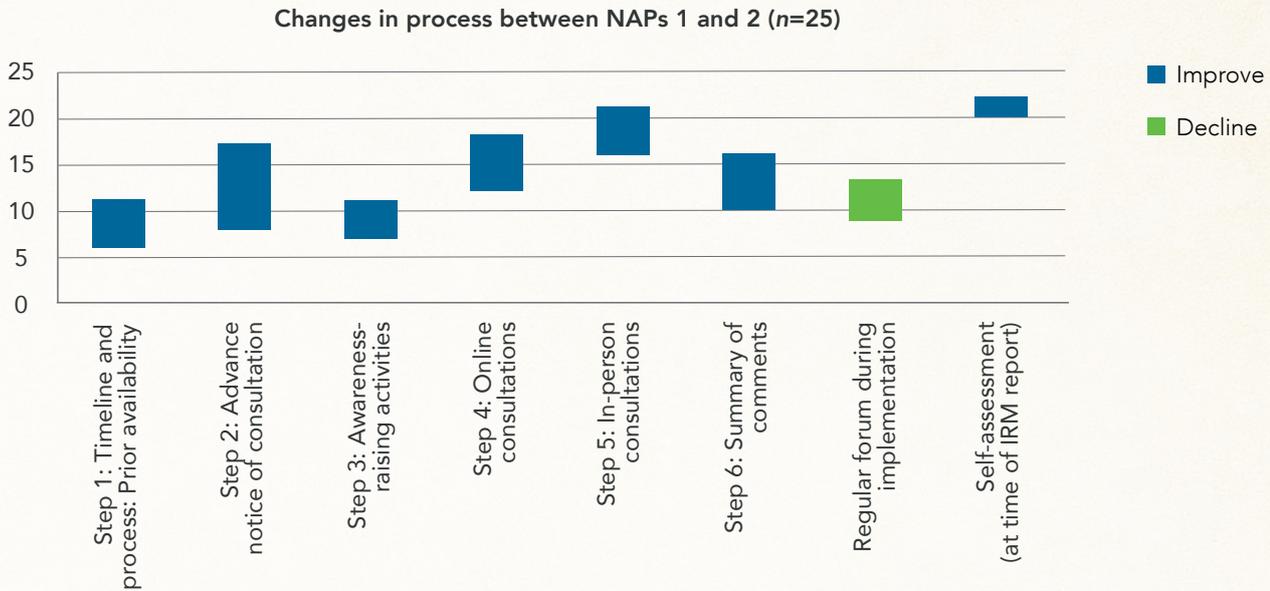
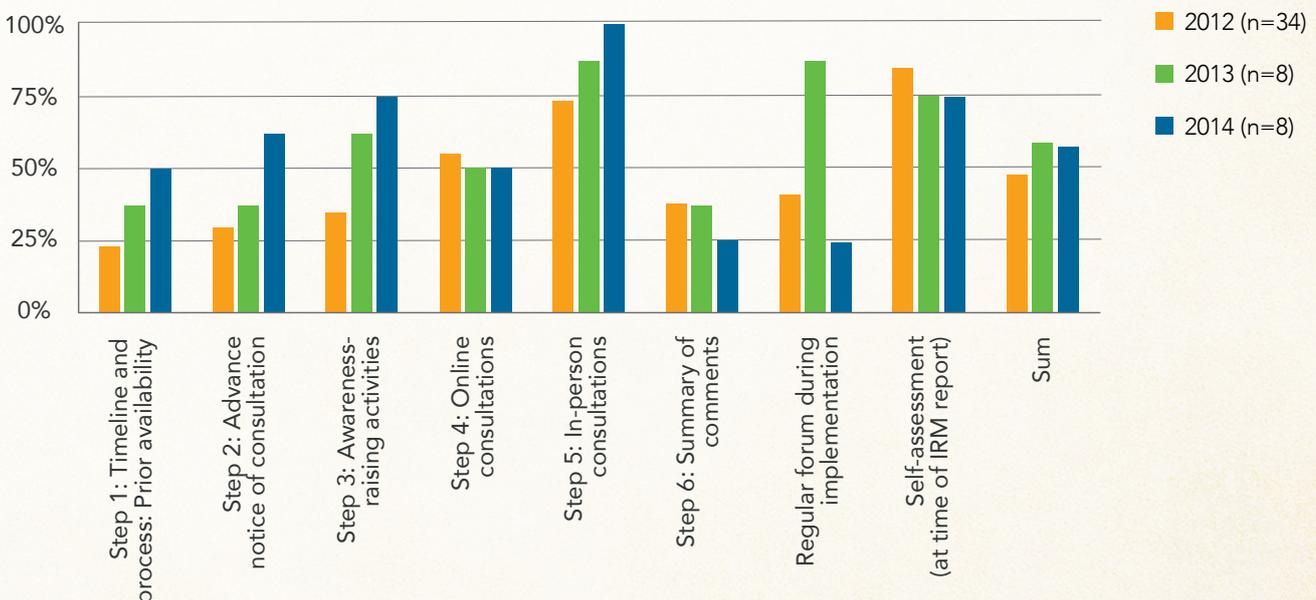


Figure 4 shows that conformity to most requirements has generally stayed the same or increased across time. Ongoing consultation during implementation has dropped. It is unclear what drove the shift, as there is some conflict between this data, anecdotal reporting from OGP countries, and other forthcoming data collected by the OGP Support Unit. It could be that the IRM’s data represents data from July 2014–June 2015, which significantly predates the data from more recent surveys and efforts by governments and civil society to institute ongoing consultation during action plans.

Figure 4: Changes in process steps over time



Beyond compliance: Toward meaningful participation

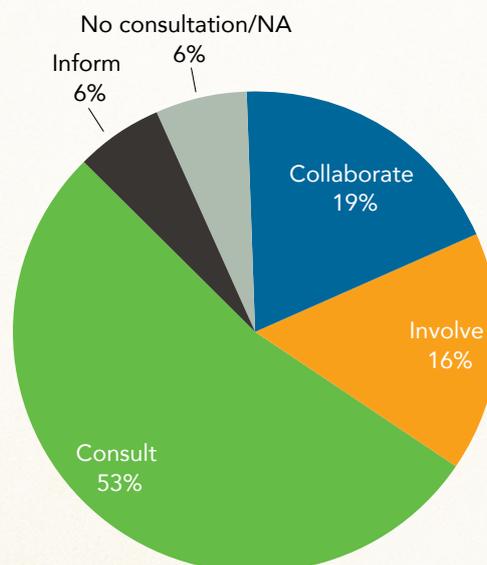
During the second half of 2016, the OGP Steering Committee is undertaking a process of ratcheting up the requirements for consultation during the OGP action plan. Some feel that there are unnecessary steps which can be de-prioritized. Still others feel that compliance with the current Articles of Governance does not guarantee meaningful participation.

Indeed, the narrative portions of the IRM suggest that in some countries governments may seek to limit participation in the action plan development process. Some reasons for doing so may be better than others. For example, budgets are a real constraint, requiring thoughtful tradeoffs. On the other hand, party affiliation of invited groups should not be a factor for participation—in the spirit of OGP, if not in the letter. In order to ensure that OGP processes were not “over-managed,” and to make a first foray into assessing the meaningfulness of participation, the IRM introduced two new indicators in 2015: “degree of potential public influence” and “open or invitation only?” Because these are new indicators, this paper cannot compare progress over time or between action plans. Rather, we can set a baseline in this paper and revisit the question in subsequent reports.

Figure 5 below shows the distribution of action plan processes across the International Association for Public Participation’s (IAP2) spectrum of public influence. The spectrum has five levels, but the top level, “empower,” (1) does not occur thus far in action plans and (2) is probably so infrequent for the types of decisions (often regulatory or legal in nature and therefore ultimately governmental) that go into OGP action plans that the IRM does not assess it. The remaining four levels range from “collaborate,” in which government officials brainstorm, craft commitment language, and monitor tasks with the public, to “inform,” in which the government does not actively seek public input. Here, the government simply pushes information at the public. (The IRM does not code for the degree to which civil society input was or was not taken into account, although it is dealt with qualitatively in the action plans.)

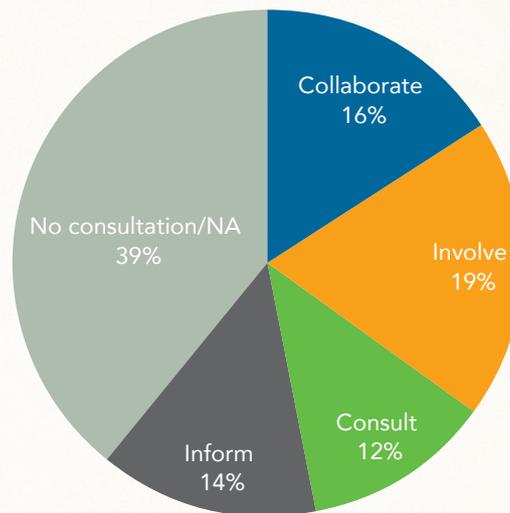
The statistics demonstrate that the majority of action planning processes (53%) ask for public input (“consult”), but there is not necessarily any sort of iterative dialogue or discussion (“involve or collaborate”). Another 12% (three countries) did not ask for any input at all (“no consultation” and “inform”). The absence of this basic level of consultation means that this subset of governments is often acting contrary to OGP principles, although the existing definition only requires them to have face-to-face meetings. It may be well past time to raise the minimum standards of listening and feedback required during the action planning process, as this high number risks the credibility of the OGP process.

Figure 5: Level of public influence over action plans during action plan development



According to the OGP Articles of Governance, governments are also required to establish ongoing means of dialogue and consultation. This may take the form of collaborative monitoring or co-implementation. Figure 6 below shows that the level of dialogue during action plan implementation was much weaker than during action plan development. Again, more than half of the action plans during implementation had no means of public input at all. In 2015, many civil society organizations and governments began building capacity for more dialogue and participatory monitoring during implementation, and this may result in raised levels of participation during implementation.

Figure 6: Level of public influence over action plans during action plan implementation



We can look beyond *how* participation occurs to *who* can participate. A number of OGP countries have been accused of selecting favorable NGOs to participate in the action planning process at the exclusion of others. In countries with strong authoritarian, corporatist traditions or fewer nonpartisan policy NGOs, there are limitations on the diversity of inputs to OGP action plans even in the best cases. At worst, OGP provides a fig leaf for the contraction or manipulation of civic space that may be occurring in parallel. As mentioned above, there are legitimate reasons for rationing participation, especially when government shares decision-making or formal advisory powers with civil society members. However, limiting participation at the level of who may attend, observe, or provide input (as opposed to vote or speak on behalf of the plan) runs counter to the spirit of OGP. While the data in Figure 7 below does not enumerate the rights of various participants (e.g., observe, comment, decide), it does show that less than two-thirds of OGP countries had open participation where any interested party could participate. Even among those who did have open processes, the IRM data is only a first step in assessing whether a broader group of civil society organizations had access to the OGP process and whether the process was inclusive. Here, more qualitative research and synthesis of existing reports is needed.

Figure 7: Openness of action plan process

| n=49 | DURING DEVELOPMENT | | DURING IMPLEMENTATION | |
|-----------------|--------------------|-----|-----------------------|-----|
| | Ct. | % | Ct. | % |
| Invitation only | 15 | 31% | 14 | 29% |
| Open | 31 | 63% | 17 | 35% |
| No consultation | 3 | 6% | 18 | 37% |

CONCLUSION: PROCEDURAL IMPROVEMENTS ARE PROMISING BUT NOT UNIVERSAL

There is clear evidence to suggest that, overall, countries complied with the minimum standards for consultation set out in the OGP Articles of Governance. However, a sizeable minority of countries continue to lag behind, and still others may formally comply with the requirements of OGP, if not the spirit. Over time, we hope that the new OGP co-creation guidelines encourage countries to collaborate and to formalize roles in the OGP process. Beyond the upward revision of participation requirements in OGP, it also requires continued investment in making sure governments and interested civil society organizations understand these requirements to move toward maximal development.

III | ACTION PLAN CONTENT AND IMPLEMENTATION

We see generally positive trends in the increasing quality of the planning process. But do we see matching trends in the formal quality of action plans? In turn, do we see improvements in implementation?

The IRM defines the quality of action plans and the commitments that comprise the bulk of those action plans along four parameters:

- **Specificity and measurability:** Commitments are rated with one of four values (none, low, medium, and high) as to the level of specificity and measurability or verifiability.
- **Relevance to OGP values:** OGP commitments must attempt to affect or use either access to information (whether proactively released or reactively released), civic space (whether affecting the environment for participation or creating specific spaces for public input), or public accountability (holding officials answerable and responsible to the public for decisions and actions). Commitments with text that does not explicitly lay out how they will affect or employ one of these open government values are marked as having “unclear” relevance.
- **Potential impact of commitments:** To assess the potential impact of a given commitment, IRM researchers assess the (1) status quo of a given policy area; (2) the goals of the given commitment; and (3) the degree to which, if implemented, the activities of the given commitment would “move the needle” in a policy area. This judgment is made independent of the clear relevance of a given commitment to open government. A commitment can have one of five ratings—worsens, none/no effect, minor, moderate, or transformative. Note that this measurement does not assess the overall importance of the policy area to national dialogue, which is treated in narrative only rather than through a ranking.
- **Implementation (and, by proxy, feasibility):** Finally, the most fundamental aspect of an IRM researcher’s task is to evaluate completion, judged by a literal and precise reading of the text of a commitment. This can receive a rating of “no progress,” “limited,” “significant,” or “complete.”

For mid-term progress reports, these four elements are synthesized into a composite rating, “Starred Commitments.” Starred commitments must meet the following criteria:

- “Medium” or “High” specificity and measurability.
- Clear relevance to an OGP value.
- “Transformative” potential impact. (Note: Prior to reports published in late 2016, both “transformative” and “moderate” potential impact commitments were eligible for star status.)
- “Significant” or “complete” progress at mid-term.

Note that beginning in 2016, the IRM began publishing end-of-term reports and introduced the variable, “Did it open government?” A forthcoming paper will summarize the results of this new measure and analyze the degree of progress in the second year of commitments.

COUNTING STARS: FROM ACTION PLAN “LEANNES” TO “MODEL” COMMITMENTS

Before assessing more granular progress, it is useful to look at the overall picture. Any assessment of OGP’s effectiveness must look at whether the action plans are delivering major reforms.

In the early days of OGP, the IRM, under the guidance of the International Experts Panel (IEP), was focused on measuring in a way that created incentives for governments to make action plans leaner and more focused, with fewer vague or irrelevant commitments (e.g., e-government and general reforms that did not open government). In order to incentivize this behavior on the part of governments and civil society collaborators, the IRM introduced the system of “starring” commitments. These commitments would be (1) specific and measurable, (2) clearly relevant to open government, (3) would have notable potential impact if implemented, and (4) were on their way to implementation. The marker of a quality action plan would be one that had a high *percentage* of starred commitments. The correct unit of analysis to evaluate “leanness” was the action plan as a whole.

One unintended consequence of this definition, however, was that many commitments that were positive, but otherwise not model commitments, received stars. (Remember that this aimed to raise the floor, not identify “model” commitments.) There was widespread—albeit understandable—misreading of the star system; many assumed that these commitments were models for other countries to emulate. Understanding that this was a common sense interpretation, the IEP changed the star system to identify model commitments. This meant tightening the “potential impact” threshold for becoming stars. Commitments now must have “transformative” potential impact in order to receive a star.

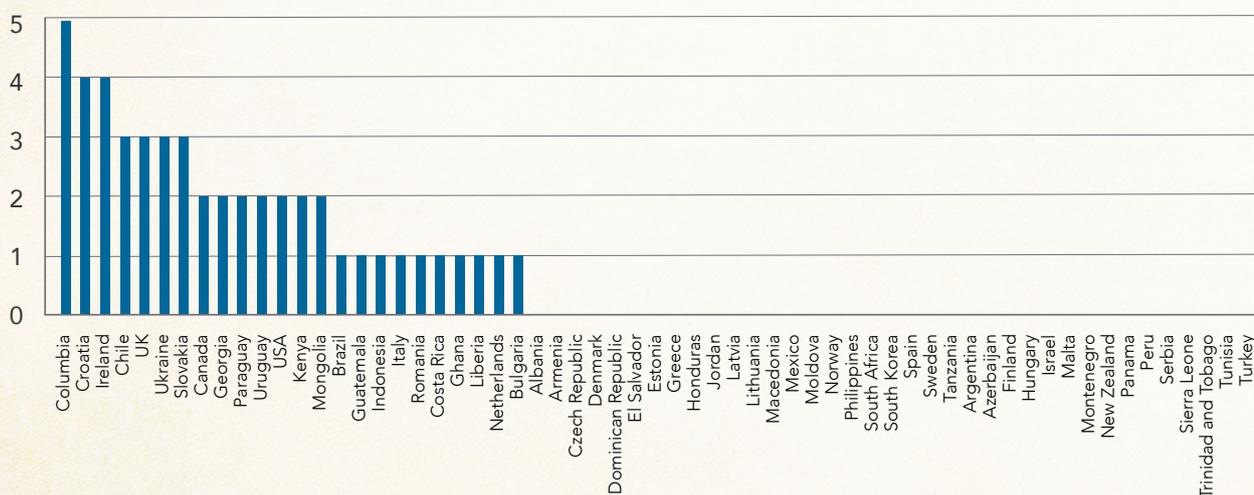
Two effects flowed from this. One immediate consequence of tightening the criteria was that the number of stars in each action plan dropped, sometimes drastically. The average dropped from 20% under the old system to less than 5% under the new system. A second consequence was that the IRM, as a whole, became much more deliberate about how and when it marked a commitment as “transformative” as opposed to “moderate” potential impact. Because of these shifts, the IRM recognized that the overall *number* of stars per action plan, rather than the percentage was a better measure of the efforts of each OGP country.

For purposes of this paper, we will look at both measures—the percentage under the old system and the number of stars under the new system—as one illuminates action plan leanness, while the other looks at OGP’s major accomplishments.

Of course, “transformative” potential impact is not the only metric to consider. There are any number of excellent commitments which should be included in action plans even though they do not fit the definition of transformative. They just do not receive stars under the new system.

Nonetheless, most would agree that the number of stars in each action plan remains too low, and too many action plans have not delivered starred commitments. Figure 8 below shows the distribution of stars per action plan in each country’s most recent action plan. (Note that this is at mid-term, and many of the countries’ star counts would presumably rise as they finish more commitments in the second half of their action plans. In other words, it is likely that the overall number of stars will be higher in the second half of the action plan.)

Figure 8: Stars in most recent mid-term progress reports



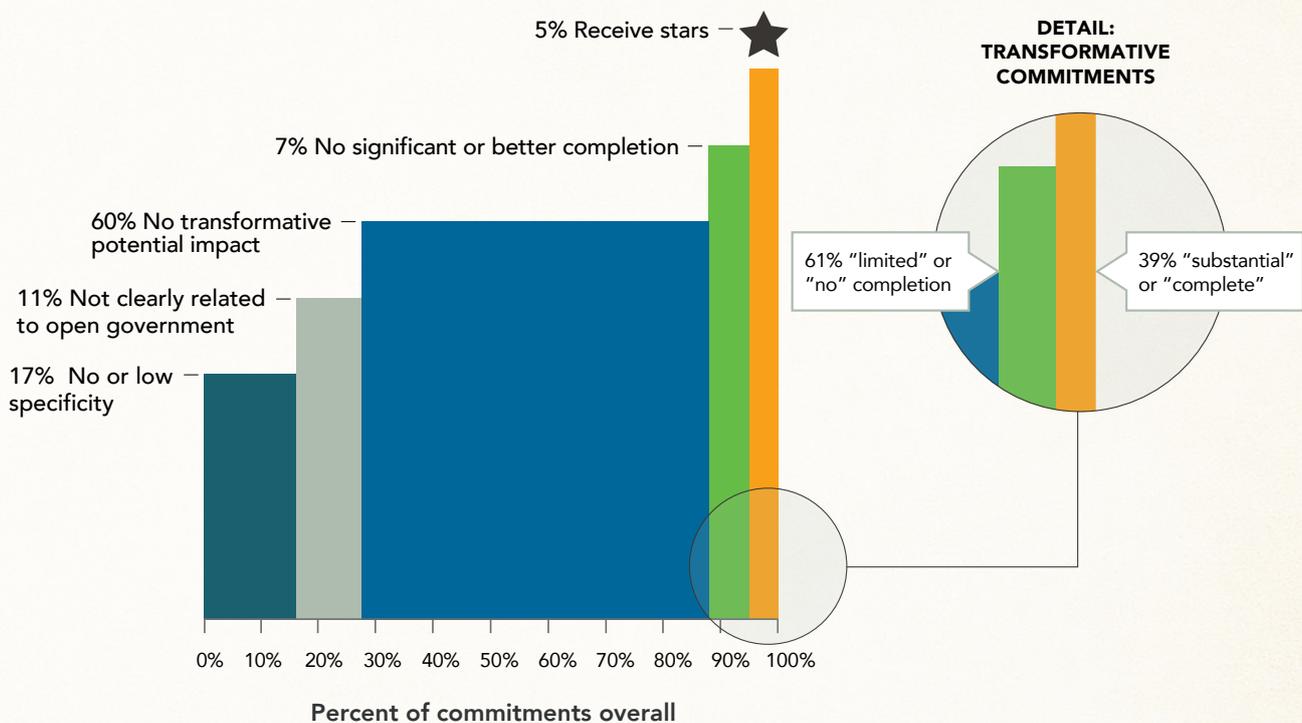
WHY AREN'T THERE MORE STARS?

A binding constraints analysis is most useful for diagnosing why we are seeing this decline. In general, by far the biggest reason commitments do not get stars is that commitments do not get high enough ambition ratings.

Figure 9 below shows the binding constraints for stars. Each of the steps is necessary to achieve the final step of getting stars. First, commitments must be specific enough to be clearly evaluated. Seventeen percent of commitments do not pass this barrier. Of the remainder, 13% are not clearly relevant to opening government (11% of the total). Out of that group, fully 83% of commitments are not rated as having a “transformative” potential impact (60% of the total). Out of the commitments that did not pass this test, 61% do not see significant progress or completion during the first year of the action plan.

Therefore, the biggest barrier to achieving starred commitments is achieving a rating of “transformative” potential impact. This is followed by implementation, which also is a major barrier. In the coming year, following a large batch of end-of-term reports, the IRM will be able to assess to what degree commitments were completed in the second year of each action plan.

Figure 9: Why commitments do not get stars



WHAT CAN STARRED COMMITMENT TELL US ABOUT THE DIRECTION OF CHANGE?

Second only to whether OGP countries are making major changes is the question of whether OGP countries are learning—namely, is the direction of change positive? In order to evaluate this across time, we retroactively applied the *new* starred commitment formula to commitments that were evaluated under the old system, and we also compared new action plans using the old system. The former is good for catching variation in reforms of “transformative” potential impact, while the latter is good for seeing variation in action plan “leanness.”

Divergence in action plan “leanness”

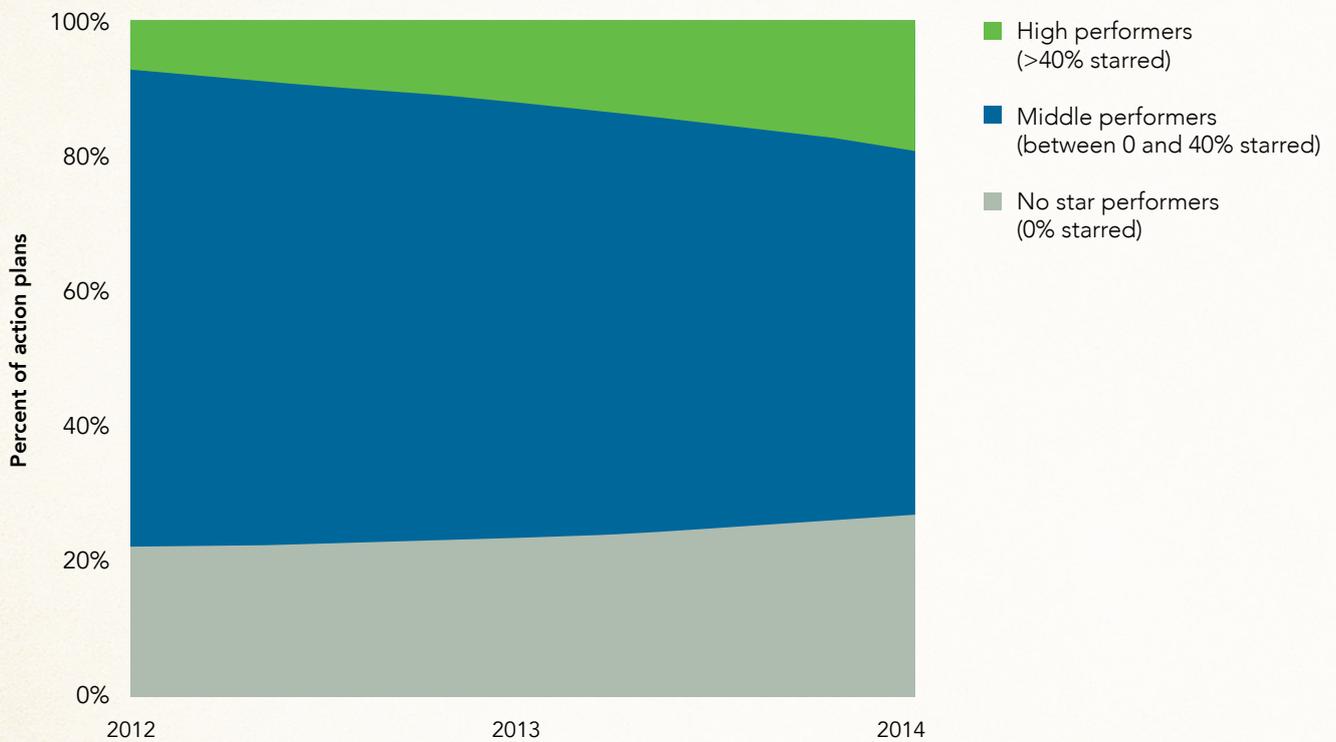
Anecdotal evidence suggests that some countries are struggling with OGP, while others are thriving. We wanted to see if this is the case, and, if so, are there longer term trends?

To assess changes, we divided OGP action plans into three categories, using the 2015 formula for the percentage of starred commitments:

- No-star countries: Those with no stars
- High performers: Those with 40% or more commitments with stars
- Middle performers: Those in between the other two categories

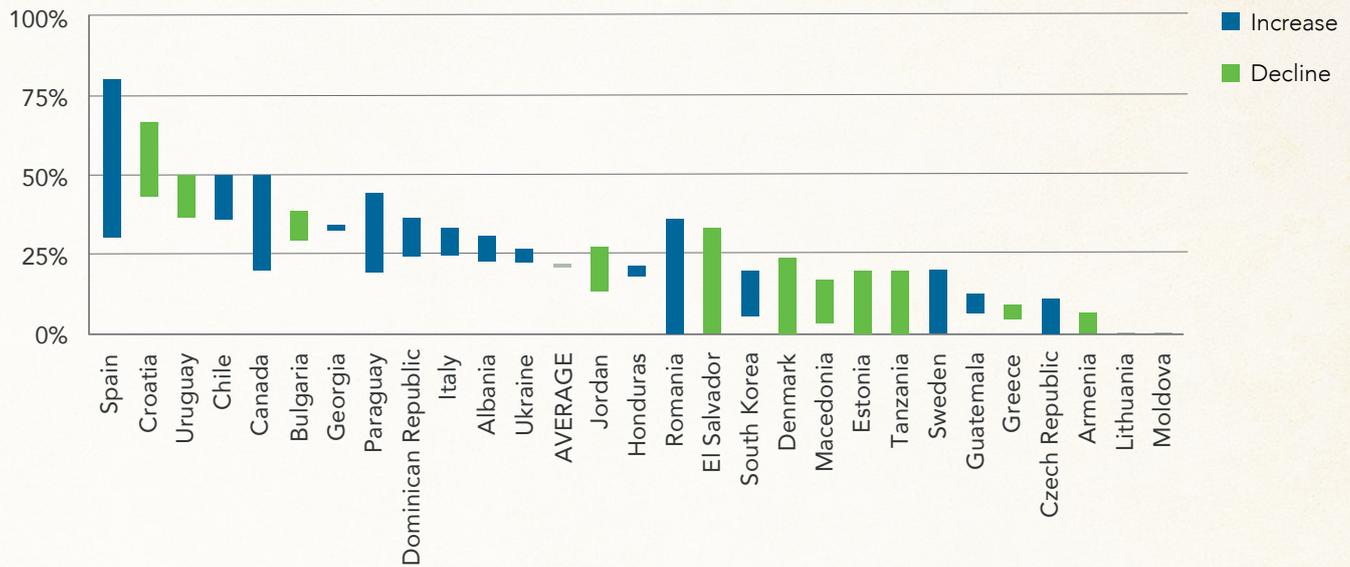
As shown in Figure 10 below, the number of high-performing countries (40% or more stars under the old formula) increased in 2014.

Figure 10. The shrinking middle



In Figure 11 below, we can see that the group of strong performers is not necessarily consistent from action plan to action plan, and some countries excel in the first action plan only to falter (at least formally) in the second. While some high performers stayed relatively strong (e.g., Uruguay and Chile) and some low-star-count countries stayed low, a large group of countries with two IRM reports moved significantly up or down. This suggests there may be some sort of mobility within OGP. In addition, some of the newest countries, such as Sierra Leone, and many of the founding countries, such as the United Kingdom, were also high performers by the starred commitments (old formula) measurement. Many of such countries are not included in the figure below as they were part of the founding cohort, so their first action plans were assessed prior to the introduction of the potential impact and starred commitment criteria.

Figure 11. Changes in percent starred commitments between NAPs 1 and 2 (old formula)



First, Figure 12 below shows the distribution of starred commitments (under the old formula) by quintile and year. The overall distribution is fairly stable with a small amount of natural variation. This shows that, generally speaking, the largest group remains low performing (<20% starred commitments) with a larger group in the middle. Between the 2012 and 2014 action plans, the chart had no major shift to the right as we had hoped. Even if we assume that IRM reports affect the way new action plans are written and implemented, there may be a natural lag; the star criteria were introduced in 2014 as well, so some of the 2014 action plans were well-formed before the IRM reports would have been read and digested. It is up to the reader to decide how quickly the chart should shift to the right.

Figure 12: Performance by year

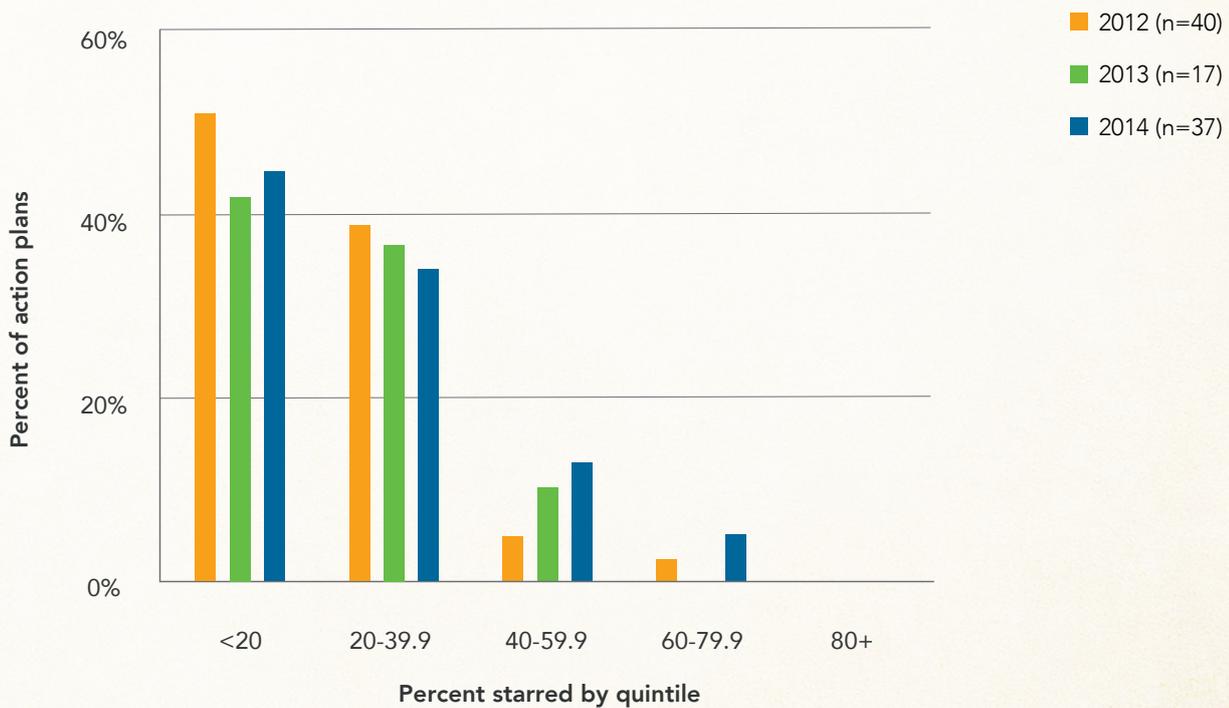


Figure 12 shows that, broadly speaking, there are only minor shifts across years, although a group of high performers seem to emerge. As these are all different countries, it is perhaps more suitable to ask whether countries improve from action plan to action plan.

Falling stars: Understanding changes across action plans

In order to understand whether and how more countries are successful in action plan development and implementation, we can use a variety of measures across different action plans. Figure 13 summarizes these measures using the retroactively applied old formula and the new star formula. Under the old formula, the percentage overall remains relatively constant (Line A). The average per action plan, however, declined, although not by a large or statistically significant amount (Line C). By contrast, using the new formula, a more drastic fall is apparent. (See Lines B and C which show declines of one-third and one-half respectively.) This is even more drastic when non-repeating countries are included.

Figure 13: Changes across first and second action plans

| Measure | | Action Plan 1 | Action Plan 2 | Increase/ decline |
|---------|---|----------------|---------------|-------------------|
| A | Starred old formula | 20% (242/1226) | 21% (142/662) | + |
| B | Starred new formula | 6% (90/1226) | 4% (31/662) | - |
| C | Average count per action plan (old formula) | 5.8 | 5.1 | - |
| D | Average count (new formula) | 1.38 | 0.7 | - |
| E | Average count only repeaters (new formula) (30 countries) | 1.9 | 0.7 | - |

A number of explanations can be given for why there is an apparent decline across action plans. We propose six possible hypotheses:

- 1. Change in calendar:** Early on, OGP countries did not have a clear calendar. Many made one-year commitments. Following the confirmation of a two-year calendar, many countries may have made longer commitments. Because IRM reports are carried out halfway through an action plan, we would see a decline in "completion" ratings without a drop in "potential impacts," if this hypothesis is true.
- 2. Changes in the star formula:** There could be a number of problems with the application of the IRM method. The first, most likely, possibility is that the change in the "potential impact" formula changed the behavior of the IRM researchers and IEP reviewers. When the IRM changed the cutoff point for stars from between "minor" and "moderate," the result was that IRM researchers were much less likely to give "transformative" impact rather than "moderate" impact. If this was the case, we should see that the percentage of the sum of commitments that are moderate or transformative stays stable, while the percentage of transformative commitments declines.
- 3. Government incentives and learning:** This hypothesis is based on the idea that commitments are less ambitious over time. On the one hand, one might say that this is because they achieved all of the major reforms early on. On the other hand, perhaps governments scaled back ambition in order to make commitment completion more feasible. At this point in OGP, this would be indistinguishable from Hypothesis 2 above. However, in a few years, we should see an increase in accountability.

4. **Inter-researcher reliability:** The IRM researchers may be more or less reliable. There may be some inter-observer reliability issues—different researchers applying a qualitative method differently or subjectively—although this should decline over time as internal quality control processes become stricter and training becomes more intensive. (Indeed, the IRM staff was very small, and the IEP was unable to review each report multiple times in the earliest days of OGP.) Further, this would not result in such a strong directional change as it would be near-conspiratorial to suggest that dozens of IRM researchers colluded to downgrade hundreds of commitment reviews. While we might see change at the country level when there is a change of researcher, we would not see across the board changes among researchers, so while there may be reasonable critiques of the IRM method of describing potential impact, they could not explain the variation.
5. **Decline in specificity:** Here we would see action plans becoming less specific and often longer.
6. **Decline in relevance:** The data would reflect more commitments that were focused on e-government or good governance, rather than open government.

The data do not support most of these hypotheses. Figure 14 below gives a summary of the median changes in the key data areas seen among the 25 countries with two action plans. It shows that there is evidence of improvement in terms of specificity, relevance, and completion. On the other hand, there are declines in ratings for potential impact regardless of the cutoff point (although it is more significant for declines in the transformative rating). This suggests either that (1) the IRM has become more rigorous in how it assesses potential impact (Hypothesis 2 above) or (2) governments have become less ambitious for whatever reason (Hypothesis 3 above).

Figure 14: Changes between median scores for NAPs 1 and 2 in key commitment attributes

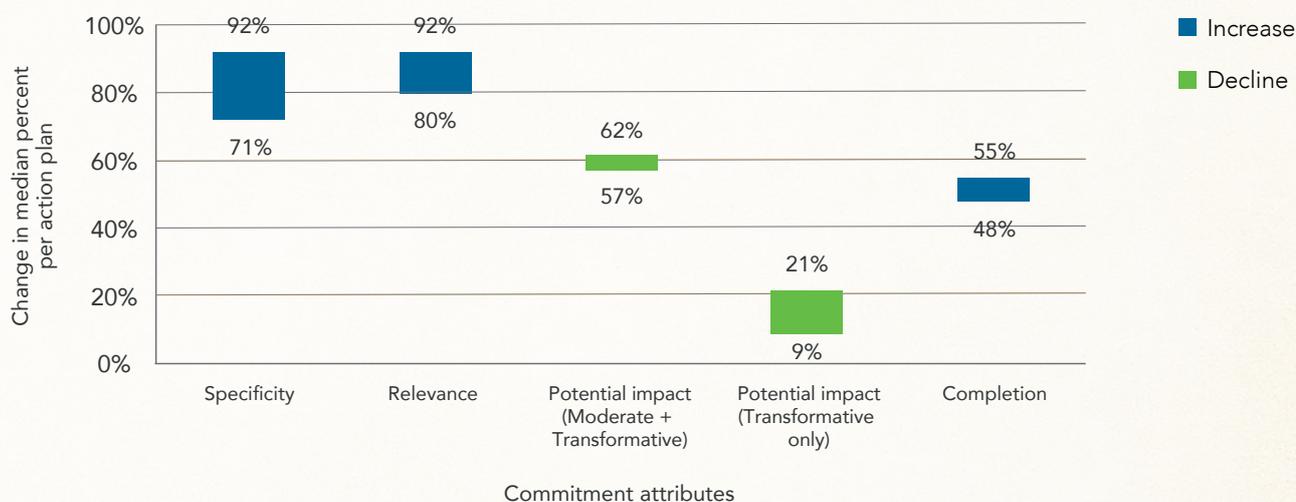


Figure 15 below compares the evidence in increase and decline with the various explanations given by the six hypotheses. Again, the pattern of evidence mainly points to the “change in star formula” and “government learning” hypotheses. There is certainly evidence that the change in the formula affected the outcomes and rankings. At this point, though, it is hard to make any assumptions about whether the “government learning” hypothesis is true. In 2017, when the IRM publishes its first end-of-term reports, we will be able to evaluate the final completion of action plans. If the rate of completion has gone up significantly, we may conclude that governments and civil society colleagues have scaled back vague, wildly ambitious action plans, like those of 2011 and 2012, to more achievable action plans.

On the other hand, if there is evidence for restrained ambition on the part of governments, this may be troubling for the OGP model (assuming it does not adequately, over the long run, balance ambition and feasibility). The model assumes that more civil society participation will result in greater ambition. If, despite the evidence of increased civil society participation in most countries, we do not see improvements over time, we might call into question some of the assumptions of the OGP model, or more reasonably whether the indicators here adequately assess the degree to which civil society participation is influential.

Figure 15: Hypotheses explaining decline in stars

| Hypothesis | | Domain response | | | | | Confirmed? |
|---|------------------------------|--|-----------|---|-----------------------------------|------------|------------|
| | | Specificity | Relevance | Pot. impact (Moderate + Transformative) | Pot. impact (Transformative only) | Completion | |
| IRM evidence | | + | + | - | - | + | |
| Expected trends in data | The calendar explanation | .. | .. | =/+ | =/+ | - | Unlikely |
| | Changes in the star formula | .. | .. | =/+ | - | .. | Possible |
| | Government incentives | .. | .. | - | - | + | Possible |
| | Inter-researcher reliability | Maintenance or increase in standard deviation of rankings over time. | | | | | Unlikely |
| | Decline in specificity | .. | .. | .. | .. | .. | Unlikely |
| | Decline in relevance | .. | .. | .. | .. | .. | Unlikely |
| Key: + Predicted increase - Predicted decrease = Remains the same .. Does not matter/affect prediction | | | | | | | |

Finally, there is little evidence of increased researcher unreliability. While there are some notable cases where the percentage of starred commitments swung wildly between researchers (e.g., El Salvador) or between action plans despite having a consistent researcher (e.g., Spain or Canada), it is hard to attribute any *directional* (up or down) change to researcher inconsistency.

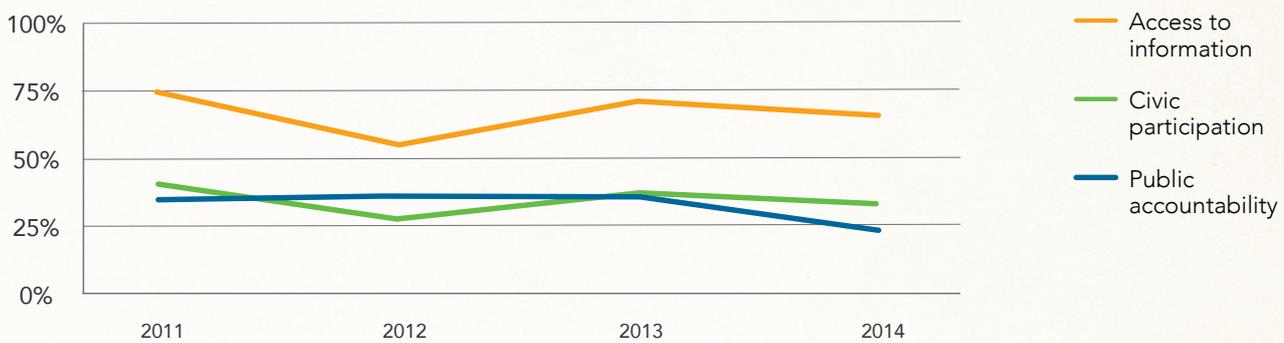
MOVING TOWARD ACCOUNTABILITY?

Those familiar with governance work will know the now commonplace refrains that “Transparency is not enough” or “Transparency and technology are the ‘low-hanging fruits’ of policy change.” This might be supported by the fact that most action plan commitments focus on transparency efforts in place of public accountability. Indeed, on average, 60% of commitments aim to improve transparency. Meanwhile, 30% of commitments focus on “public accountability.” (The labels are not, however, mutually exclusive.) This is shown by Figure 16 below.

Some have argued that over time the focus on transparency would shift toward a greater focus on accountability as the ministries, departments, and agencies, after having focused on releasing information, would move toward giving citizens the means of providing feedback or triggering accountability measures. This is not in evidence in Figure 16, which shows a drop in the number of commitments related to accountability.

There are at least four non-exclusive reasons for why this might be. First, most undoubtedly, the IRM has become stricter in evaluating commitments. In order to be tagged “accountability,” there must be a public-facing element to the accountability; a commitment, however worthy in its ambition and implementation, cannot be considered “open government” if it is secret and completely internal to government. Second, the IRM now evaluates the commitment text closely. It must articulate how members of the public can trigger a review for decisions or actions taken by the government. While these means do not need to be formal—they can, in some circumstances, follow the “Yelp” model of public feedback (see Fox and Peixoto’s 2015 paper, “When Does ICT-Enabled Citizen Voice Lead to Government Responsiveness?”)—they should at least articulate the citizen feedback loop and how the government would act on public input. This standard was not in place when evaluating the 2011 and 2012 action plans, so it is likely that the observed difference, to some extent, is attributable to a change in the IRM method.

Figure 16: OGP values in action plans by year



Yet the bigger picture is that the trajectory for accountability in Figure 16 remains largely unchanged, at half of the amount of transparency. So a second hypothesis to explain the lack of accountability commitments is that governments remain resistant. This resistance may be due to technical or political difficulties in developing and implementing such commitments, as there are often winners and losers in accountability measures.

Anecdotally, within action plans, there is a tendency to not articulate what feedback loops are in place, formal or informal. Indeed, some have argued that in some languages, such as Spanish, the word for accountability connotes a “public reading of the accounts,” rather than an actual feedback loop. Since no one would argue that the Spanish language (or any other language), regardless of existing vocabulary, would render citizens incapable of holding officials publicly responsible for their decisions and actions, there is no reason to imagine why action plans can’t articulate this loop. Assuming we believe this hypothesis is true, a solution could be to educate OGP stakeholders on how to better design commitments which have an express feedback loop element.

A third hypothesis is that agencies would develop accountability commitments but are unable to because the agencies in charge change so frequently. Currently, the IRM does not collect the data to map the full amount of agency turnover but will do so over the course of the next several years.

Finally, there is a reasonable argument that two years is too little time to see a radical shift in priorities or understanding of how open government can be used to make officials more accountable. A shift should be evident, following this logic, after a few more action plans.

IV | TYING IT TOGETHER: ACTION PLANS AND NATIONAL-LEVEL CHANGE

The previous sections showed evidence that OGP action plans are creating change that is generally positive.

While there are countries that are not taking full advantage of OGP, others are capitalizing on the OGP platform and improving government-civil society dialogue. Indeed, there is little reason to believe that these improvements will not continue as the number of civil society organizations and ministries involved increases, as governments formalize their processes for OGP, and as OGP Support Unit staff and partners increase bilateral and multilateral support.

THE RELATIONSHIP BETWEEN BETTER PARTICIPATION AND IMPLEMENTATION

The 2014 IRM Technical Paper turned up a weak, but intriguing, correlation between rates of completion and the level of participation and no relationship with the percentage or number of starred commitments. In this paper, we have seen a general rise in participation and a small rise in completion, with a diminishment in potential impact. But this is at the global aggregate level. Is there a relationship at the level of the individual action plan?

In order to analyze these relationships around individual action plans, we looked at three dependent variables—specifically those which are “outputs” of OGP:

- Percentage of complete commitments;
- Percentage of substantial and complete commitments; and
- Starred commitments

We identified independent variables that might have some influence on the success of an action plan.

- How many of the six steps for consultation during action plan development were followed? (Steps followed)
- What level of influence did the public have over development of the action plan?
 - Citizens were asked their opinions (Level of consultation (“Consult” or better))
 - Citizens received feedback on inputs (Level of consultation (“Involve” or better))
- Was there ongoing participation during action plan implementation in the country? (Ongoing participation)
- Was there a change in president or prime minister during the period under assessment by the IRM? (Change in executive)
- Was there a change in the agency or ministry in charge of implementing OGP. (Multiple arrangements)

Figure 17 below shows the relationships (or lack thereof) between each of the dependent and independent variables. Despite a larger dataset than the 2014 IRM technical paper, the same tests continue to show no particular statistical significance. The two exceptions are between “substantial or complete level of completion”

and starred commitments with multiple governance arrangements. If a government had a change of governance structure, it was less likely to implement its commitments. An r^2 of 5% and 7% respectively suggests that, while this is not a major factor in implementation, it can explain a non-negligible amount of variation.

Figure 17: Regression analysis of outcome and process variables

| | | Steps followed | Level of consultation ("Consult" or better) | Level of consultation ("Involve" or better) | Ongoing participation | Change in executive | Multiple arrangements |
|---------------------------------------|-----------|----------------|---|---|-----------------------|---------------------|-----------------------|
| Completion (complete only) | r^2 | 0.004 | 0.000 | 0.040 | 0.007 | 0.010 | 0.003 |
| | P | 0.576 | 0.943 | 0.169 | 0.481 | 0.387 | 0.627 |
| | direction | + | - | + | + | - | - |
| Completion (substantial and complete) | r^2 | 0.020 | 0.005 | 0.043 | 0.009 | 0.001 | 0.050 |
| | P | 0.194 | 0.647 | 0.151 | 0.415 | 0.742 | 0.051* |
| | direction | + | + | + | + | - | - |
| Stars | r^2 | 0.031 | 0.038 | 0.011 | 0.002 | 0.000 | 0.042 |
| | P | 0.107 | 0.177 | 0.475 | 0.673 | 0.940 | 0.074* |
| | direction | + | + | + | - | + | - |

The implications are several. First, the relationship between better participation and better implementation may be multifactorial and would not easily be picked up in a single-variable correlation analysis. Second, there is no evidence to suggest that a change in executive is necessarily a negative factor in implementation. Third, there is some evidence to suggest that institutional stability (and perhaps, by proxy, institutional learning) might contribute to rates of completion and significant (starred) reforms.

BEYOND THE ACTION PLAN: NATIONAL-LEVEL CHANGE

Beyond the relationship between the participatory process and commitments, OGP will need to be aware of what the aid effectiveness experts refer to as the "micro-macro paradox." Namely, why don't the foreign direct investments and international development activity add up to the sum of their parts? (For the longer discussion of this, see Mosely's "Aid Effectiveness: the micro-macro paradox" 1986 and Picciotto, 'Development Effectiveness: An Evaluation Perspective' 2009)? Specifically, if so many individual country assistance programs receive positive evaluations, why do so many aid recipients continue to suffer from economic stagnation?

Extrapolated to OGP, if so many action plans are evaluated positively (which they are not necessarily), then do we see evidence of country-wide change? As with any intervention, there are tradeoffs. Money spent on OGP—while not a lot by international standards—could be spent in other governance interventions. It is possible—although it is not our current opinion—that the action plan as part of the move to more open government might not be the best fit. Some have argued that it is too short or too managerial to have larger impacts. Alternately, and more immediately, action plans do not often deal with system-wide changes that would aid in promoting open government.

So, let us assume that performance in OGP action plans is an independent variable. The dependent variables would be open government and then, according to this logic, greater public integrity, citizen voice in decision making, and civic space.

In an attempt to be accountable for overall outcomes, OGP staff tracks a number of third-party indicators which might show the overall progress of OGP countries in key areas around OGP values. The indicators include and go beyond the official OGP eligibility requirements. These include (with abbreviations and years in parentheses):

- **Access to information**
 - Open Budget Index (OBI) (2010, 2012, 2015)
 - OGP's scoring of right2info's Right to Information law (ATI rating) (2012, 2015)
 - Open data barometer (ODB scaled score) (2013, 2015)
- **Civic participation**
 - Economist Intelligence Unit Civil Liberties Score (EIU CL) (2012, 2013, 2014)
 - The World Bank World Governance Indicators, Voice and Accountability Index (WGI V&A) (2012, 2013, 2014)
- **Accountability**
 - World Justice Project Rule of Law Index (RoLI) (2012–13, 2014, 2015)
 - Asset Disclosure (AD) scores from OGP Eligibility Requirements (2012, 2015)

Yet, some caution is warranted in attempting to link OGP action plan results with changes at the national scale, at least as measured by third-party indicators. There are a number of factors which make such analysis not only difficult but also too much of a stretch in logic. Some of the reasons include:

- **Time frame:** The action plans assessed in this paper examine the commitments written in 2011–2014—two and three years after OGP was founded. Similarly, many of the datasets, even those published in 2015 are based on observations from 2014. It may be unrealistic to think that OGP could contribute to nationwide change in such a short period of time.
- **Self-selection:** Prior to joining, OGP countries already had significantly higher scores across all indicators. Those countries which were *eligible* but did not join OGP had lower scores than participating countries. Thus, there is some self-selection going on before joining OGP, suggesting that countries which joined already had demonstrated greater commitment and had more robust legal frameworks in place to promote open government. This can be seen in Figures 18.1 18.2 and 18.3 below.
- **Proving the counterfactual:** As with any such exercise, being able to prove what would have happened without the OGP action plan is difficult, especially since there is self-selection in OGP membership to begin with. There may be a decline in democratic values and other values associated with openness worldwide (although this is hotly debated), but not among current or most potential OGP countries. (Again, see Figure 18.2 below with the Economist Intelligence Unit Civil Liberties score.)
- **Scope of action plans:** The scope of OGP action plans is not congruent with the major OGP values in many countries. Where OGP countries do have deficits, many countries are not addressing many key reforms identified by lower scores. Of course, what is in action plans may only reflect some portion of the total reforms that an OGP country is undertaking. When this is the case, however, how OGP might affect larger indicators becomes more circuitous or diffuse.

- **Different value propositions:** Finally, it does not follow that the likely results or aims of OGP will be uniform across participating countries. The value of participating may be very different for an OECD member with very strong institutions than it is for a country acceding to the European Union, or a heavily indebted poor country. Measuring success solely on macro-level indicators may miss the actual value that OGP can bring.

When we begin looking at specific indicators, the picture is decidedly mixed. Among some indicators, there seems to be no correlation between OGP participation (civil liberties and open budgets, for example). In others there seems to be a promising correlation (open data). Figures 18.1, 18.2, and 18.3 show a subset of third-party indicators most closely aligned with OGP values. They are by no means exhaustive and are broken apart only for ease of reference and reading.

Figure 18.1 below shows some of the difficulty in assessing any change at all. Let us take the three points of data from the Open Budget Index as an example. Between the three categories of countries, we saw a clear directional shift. More governments worldwide are taking steps to make the budget process more open and participatory. But while the changes overall are in a positive direction, the most movement is seen in countries that are ineligible for OGP, both as a raw score and as a percentage of improvement over time. This suggests that either (1) it is easiest to make some of the incremental changes when one is at the bottom—to grasp at “low-hanging fruit”—or (2) OGP, the International Budget Partnership, and other international “standards” actually incentivize improvements in budget openness. As the IRM does not gather data on non-participating countries (let alone their motivations for improving budgets), it is outside of the scope of this paper to examine why the changes at the bottom end of the scale are improving. Within the group of participating countries, we see the changes in the OBI score to be limited, and there is little variation within the group (no countries seem to go way up or way down year on year). A future paper could look at those few cases which go up or down and correlate these to the budgetary openness commitments, but that is also beyond the scope of this paper.

If we look at other indicators, the idea that joining OGP incentivizes higher performance does not hold up. For example, among non-participating countries, on average, no new laws were passed. Were we to assume that OGP participation is a major incentive, we should see an improvement among ineligible countries. We do not, however. In fact, we see the only aggregate increase among eligible countries who chose not to be part of OGP, which flies in the face of the notion that OGP is a major motivator to improve access to information laws. Of course, perhaps a better question is whether implementation of access to information laws, where they exist, is better than passage of laws in OGP countries. This, too, is beyond the scope of this paper.

The Open Data Barometer, however, shows an interesting relationship. Here, OGP countries actually started lower than non-participating countries but improved more over time. To what extent this is due to the support of OGP is unclear. A future paper may look at the difference of means between OGP countries and the non-OGP countries in terms of how much they improved.

Figure 18.1: Open budget, access to information, and open data

| | | OBI (2010) | OBI (2012) | OBI (2015) | ATI (2012) | ATI (2015) | ODB (2013) | ODB (2014) | ODB (2015) |
|---------|-------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Scores | Participating | 56.9 | 57.0 | 58.6 | 3.9 | 3.8 | 33.9 | 39.9 | 37.3 |
| | Not participating | 41.9 | 45.8 | 47.4 | 3.3 | 3.6 | 34.6 | 34.1 | 30.8 |
| | Ineligible | 19.4 | 19.9 | 25.2 | 1.6 | 1.5 | 30.5 | 32.1 | 31.2 |
| | Average | 42.2 | 42.7 | 45.4 | 3.0 | 3.0 | 32.5 | 34.9 | 33.0 |
| Changes | Participating | | 0.1 | 1.6 | | 0.0 | | 5.9 | -2.6 |
| | Not participating | | 3.9 | 1.6 | | 0.4 | | -0.6 | -3.3 |
| | Ineligible | | 0.6 | 5.2 | | 0.0 | | 1.5 | -0.9 |
| | Average | | 0.5 | 2.7 | | 0.0 | | 2.4 | -2.0 |

Explanation: **Bold** means decline. Some declines are minor and not visible due to rounding.

Figure 18.2 covers broader operating environment indicators, including civil liberties and the World Bank Voice and Accountability indicators. Neither of these is particularly elastic, which either shows that claims of deteriorating democracy are overblown or shows that the indicators are not sensitive enough to pick up major changes in the national-level operating space. While there is a very minor decline (when rounding) for EIU civil liberties scores, it is uniform across categories and cannot be correlated with OGP participation. The WGI Voice and Accountability indicators show almost no variation. The significance of this non-finding is that these scores are not likely to change much—at least in the short term—as a result of many government actions, whether part of OGP or not. This suggests one of two things. Either new, more sensitive indicator sets could be used (e.g., the CIVICUS Civic Space Monitor or ICNL’s Civic Freedom Monitor), or OGP stakeholders should temper expectations that OGP will create these kinds of changes in the short term.

Figure 18.2: Civil liberties and voice and accountability

| | | EIU CL (2012) | EIU CL (2013) | EIU CL (2014) | WGI V&A (2012) | WGI, V&A (2013) | WGI, V&A (2014) |
|---------|-------------------|------------------|------------------|------------------|-------------------|--------------------|--------------------|
| Scores | Participating | 7.9 | 7.9 | 7.9 | 0.4 | 0.4 | 0.5 |
| | Not participating | 6.8 | 6.7 | 6.7 | 0.0 | 0.1 | 0.1 |
| | Ineligible | 4.1 | 4.1 | 4.1 | -0.5 | -0.5 | -0.4 |
| | Average | 6.2 | 6.2 | 6.2 | -0.1 | -0.1 | 0.0 |
| Changes | Participating | | 0.0 | 0.0 | | 0.0 | 0.0 |
| | Not participating | | 0.0 | 0.0 | | 0.0 | 0.0 |
| | Ineligible | | 0.0 | 0.0 | | 0.0 | 0.0 |
| | Average | | 0.0 | 0.0 | | 0.0 | 0.0 |

In terms of accountability for public officials, the Asset Disclosure Index and the American Bar Association’s Rule of Law Index also show nearly no variation. Interestingly, the asset disclosure indicators should show the most variation, as this is an indicator that strongly correlates to OGP participation (as it is an eligibility requirement). A next step would be to look at whether the asset disclosure laws and regulations were put in place right before or right after joining OGP.

Figure 18.3: Asset disclosure and rule of law

| | | AD (2012) | AD (2015) | RoLI (2012–13) | RoLI (2014) | RoLI (2015) |
|---------|-------------------|-----------|-----------|----------------|-------------|-------------|
| Scores | Participating | 3.3 | 3.3 | 0.6 | 0.6 | 0.6 |
| | Not participating | 3.5 | 3.5 | 0.6 | 0.5 | 0.6 |
| | Ineligible | 1.7 | 1.7 | 0.5 | 0.5 | 0.5 |
| | Average | 2.7 | 2.7 | 0.6 | 0.6 | 0.6 |
| Changes | Participating | | 0.0 | | 0.0 | 0.0 |
| | Not participating | | 0.0 | | 0.0 | 0.0 |
| | Ineligible | | 0.0 | | 0.0 | 0.0 |
| | Average | | 0.0 | | 0.0 | 0.0 |

In conclusion, there is not strong evidence to support wide variation in OGP beyond correlation with open data. As next steps (in another paper perhaps), one could extend the analysis and trends further by:

1. Analyzing the data at the national level, rather than in the aggregate.
2. Differentiating between countries after they have submitted their first action plan and their second.
3. Carrying out difference of means test (e.g., *chi-square*) between the different categories.

V | OPERATIONAL RECOMMENDATIONS

The findings from this paper point to a set of key recommendations and supports for the Support Unit, the IRM, and the OGP Steering Committee to improve the performance of OGP countries.

Improving the OGP process

- The Steering Committee should endorse the proposed guidelines on participation in OGP action plans, which support iterative, meaningful dialogue and maximize diverse views. Following agreement on new guidance, the Support Unit will need to support governments and civil society to adjust to revised expectations.
- The IRM will need to continue to track public influence on action plans and identify where the planning, implementation, and evaluation process is open and meaningful. The current method is limited to assessing only whether consultations are open and the type of input used, but not how civil society input into the action plan influences the final result.
- National governments can better support OGP points of contact to ensure that they are adequately resourced and stable by creating a clear incentive structure (e.g., performance reviews, job descriptions, and promotional tracks) to perform well as OGP coordinators. There is some evidence to suggest that maintaining a consistent administrative structure within the bureaucracy is more important than executive-level support. The Support Unit can work to ensure that, as much as possible, the institutions that are responsible for OGP are adequately staffed, funded, and resilient across administrations and through action plans.
- Additional research is still needed to identify under which conditions OGP capitalizes on national political dynamics in order to better promote open government.

Improving OGP commitments

- The Steering Committee, the Support Unit, donors, and multilateral partners should focus energy on two particular areas: (1) increasing and clarifying the ambition of commitments and (2) improving completion rates, especially among transformative commitments. Means of achieving this could include:
 - The Support Unit should revise templates for action plans that encourage governments to adopt a problem-oriented approach to commitment design. This should lay out the *real world* social, political, economic, or environmental problem to be tackled and how the commitment will work to improve upon the status quo.
 - The Support Unit should strongly consider strengthening the early review process of commitments to improve clarity and ambition. This could be by providing additional technical recommendations or by publishing these early reviews.
 - Once transformative or promising commitments are identified through an early review process, the Support Unit can work to direct collaborators (peers, experts, and donors) to support those transformative commitments and ensure a higher rate of completion. This will require the development of a better database of transformative commitments.

Tracking impacts

- For now, it is not possible to link OGP results with nationwide changes. Impact evaluations should focus on evolving practice—both good and bad—within OGP. These evaluations will be, by necessity, small-n studies and individual cases. Of course, such evaluations may be generalizable, but care should be taken not to over-interpret the information.
- At the same time, future evaluations of OGP success will need to be sensitive to the different ways that OGP stakeholders judge success. OGP countries have different reasons for participating, and, while some stakeholders may look for nationwide shifts, until action plans are aligned to promote such shifts, connections between the action plan process and national outcomes will be diffuse at best. This should be separated from more formal assessments where OGP's value can be weighed against what *could* be accomplished in a given country and in a given political context.



Independent Reporting Mechanism
Open Government Partnership
c/o OpenGovHub
1110 Vermont Ave, NW
Suite 500
Washington, DC 20005

