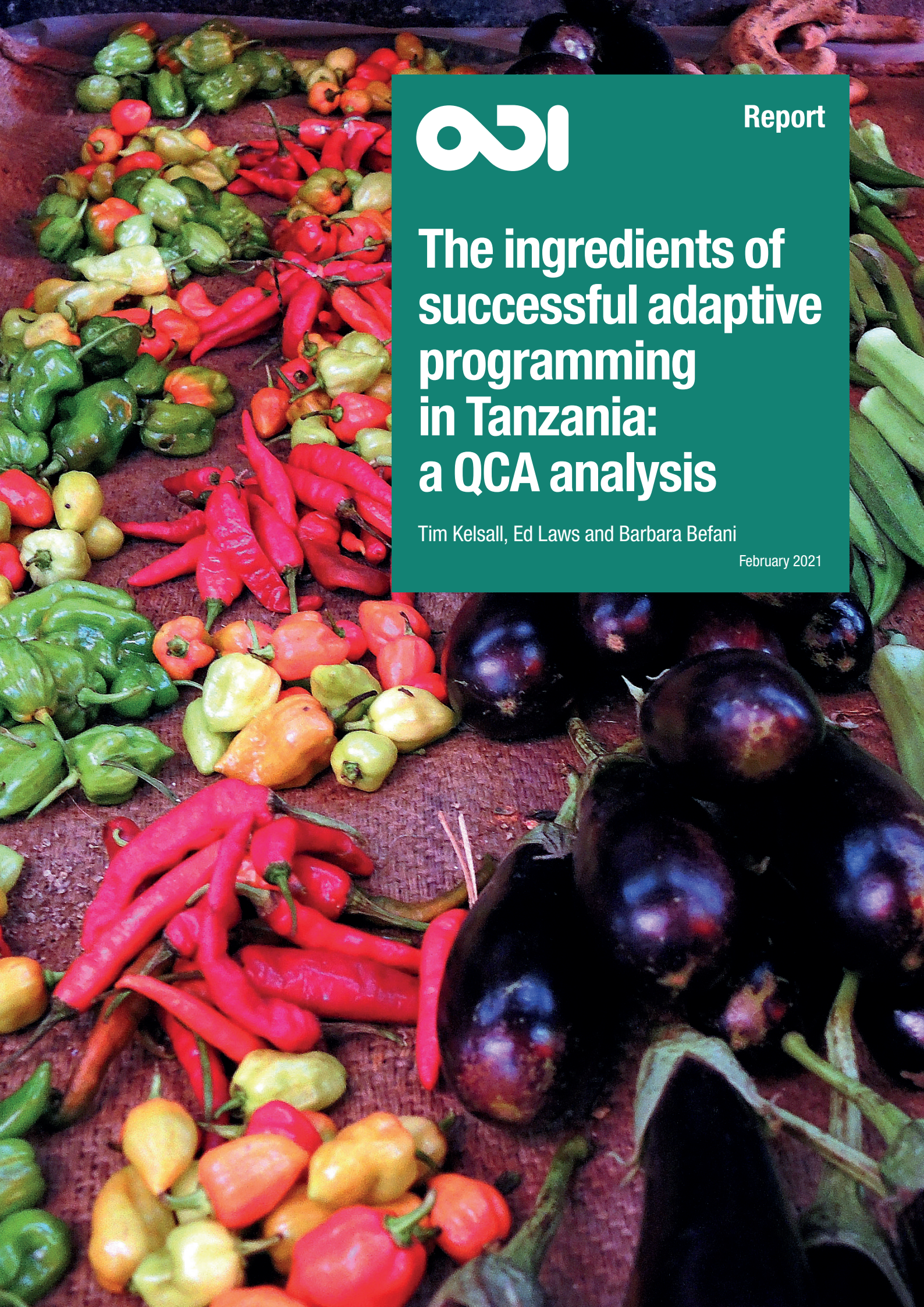


Report

The ingredients of successful adaptive programming in Tanzania: a QCA analysis

Tim Kelsall, Ed Laws and Barbara Befani

February 2021

Cover photo: Fresh produce for sale at a Tanzanian market. Güldem Üstün/Flickr.

Acknowledgements

This report is part of ODI's research and learning role within the UK Foreign, Commonwealth & Development Office (FCDO) and Irish Aid's Institutions 4 Inclusive Development (I4ID) Tanzania programme consortium. We thank all the I4ID staff who generously helped us with assembling our data, in particular Julie Adkins, Anna Bwanac, Maria Lampe and Gloria Sikustahili, with whom we met several times to discuss and fine-tune the approach; and Simon Milligan and Kathryn Nwajiaku-Dahou, who also provided helpful advice. We thank our peer reviewers, Tiina Pasanen and Gareth Williams, and our communications team at ODI, Paola Abis, Matthew Foley and Sara Hussain.

Contents

List of boxes, tables and figures

Boxes

Tables

Figures

Acronyms

CIG	Centre for Inclusive Growth (Nepal)
CSO	civil society organisation
DDD	doing development differently
FCDO	Foreign, Commonwealth & Development Office (UK)
GESI	Gender and Social Inclusion
I4ID	Institutions for Inclusive Development
IE	inclusive education
MEL	monitoring, evaluation and learning
MHM	Menstrual Health Management
MoEST	Ministry of Education, Science and Technology (Tanzania)
PDIA	problem-driven iterative adaptation
PEA	political economy analysis
PO-RALG	Ministry of Regional and Local Government (Tanzania)
QCA	Qualitative Comparative Analysis
RIFO	Regional Investment Facilitation Office
SF	Sunflower
SMEs	small and medium enterprises
SUWASA	Singida Urban Water Supply Authority (Tanzania)
SWM	Solid Waste Management
TSL	Tanzanian sign language
TWP	thinking and working politically
UNCTAD	United Nations Conference on Trade and Development
USD	Urban Spatial Development
UW	Urban Water
UWV	Urban Women Vendors

Executive summary

'Adaptive management' and 'politically smart programming' are increasingly popular ideas in development. They capture an ambition to programme in ways that are more flexible and experimental, and which respond to and capitalise on political dynamics and incentives. Over the last 10 to 15 years there has been a notable increase in aid programmes that explicitly reference these terms, or similar ideas such as 'doing development differently' (DDD) and problem-driven iterative adaptation (PDIA).

There is a growing consensus that interventions are more likely to make a positive difference in highly complex situations if they adopt these principles and methods. However, according to recent reviews of the literature on thinking and working politically (TWP) and adaptive management, much of the evidence used so far to support these approaches is anecdotal, does not meet high standards for robustness, is not comparative and draws on a small number of self-selected, relatively well-known success stories.

To generate lessons to help deepen the evidence base, this report uses Qualitative Comparative Analysis (QCA) to explore the most important ingredients of success in the Institutions for Inclusive Development (I4ID) Programme – an adaptive, issue-based governance initiative in Tanzania.

Background to I4ID

I4ID was a five-year programme funded by the UK Foreign, Commonwealth & Development Office (FCDO) and Irish Aid designed to promote inclusive development and strengthen democratic institutions in Tanzania. Beginning work in 2016, it gradually built up a diverse portfolio of workstreams, ranging from urban development to menstrual health. Despite early closure, I4ID can claim some significant breakthroughs on previously intractable or neglected issues.

Methodology

QCA uses set-theoretic logic to synthesise data and identify the patterns behind successful outcomes, including whether particular ingredients or conditions, or combinations of ingredients or conditions, are necessary or sufficient, or merely important, for producing the desired outcomes.

We analysed 40 potential ingredients of success, grouped under broad categories such as Politics, Design and Actors. We then compared the success of different combinations of these broad categories and their sub-components across four development outcomes. We then applied some broad rules of thumb to identify those combinations that might reasonably be judged to be 'core ingredients' for programmes like I4ID, in places like Tanzania. We looked for combinations, or models, which covered a high percentage of successful cases across I4ID's outcomes, or covered a smaller percentage but with high consistency.

Looking across all the results, we present the core ingredients in the form of a 'recipe' for successful adaptive governance programming (Figure 1).

Figure 1 The core ingredients of adaptive programming

STEP 1 Assemble your team	STEP 2 Identify a problem	STEP 3 Design your intervention	STEP 4 Test and learn	STEP 5 Spend strategically and stay flexible
• Enlist locally based staff, preferably with strong local networks or membership in a relevant national organisation. • Add a team leader or another senior staff member with a track record of successfully delivering an adaptive programme.	• Look for an issue that is already receiving a high level of political attention. • Use a combination of light-touch political economy and light-touch systems mapping and analysis to deepen your understanding of the problem and plot possible solutions. Apply more rigorous analysis as appropriate. • Formulate a loose theory of change or set of hypotheses about how change might happen.	• If possible, enlist positive deviants and/or leverage external best practice to start generating possible solutions. • Combine with human-centred design principles, particularly when trying to create more inclusive decision-making, planning or policy processes. • Have a credible plan in advance for taking interventions to scale.	• Test one approach or potential solution to your problem, gather feedback and adapt as necessary. • Enlist the support of implementing agencies and non-state actors, and – especially if you anticipate or encounter organised opposition – high-level political actors. • Take care to also maintain the support of your funding organisation(s). • At regular intervals, review your progress using light-touch political analysis tools, perhaps with more in-depth analysis at critical junctures.	• As your intervention matures, use programme funds in a strategic way. You may look for solutions through brokering and convening, but be prepared to provide funds for pilots, small-scale infrastructure, technical assistance or core funding. • As needs arise, and depending on your problem, you may also need to build capacity for innovative behaviour or innovative technology; alternatively, it may be sufficient to improvise on existing solutions. Maintain flexibility to be able to respond appropriately.

What does this add to the evidence base?

Anyone familiar with the literature on adaptive programming and on thinking and working politically is unlikely to be surprised by this recipe. The value of this study consists partly in the additional rigour and transparency with which we have demonstrated its relevance. We believe we have established 'proof of concept' that this recipe works.

However, as we outline in the broader story, we also think our study has some less obvious lessons to add to the conversation on these approaches – albeit with some important caveats, on which we elaborate in the main report.

Perhaps most notably, I4ID demonstrated that it can achieve success with a somewhat loose and informal approach to political economy analysis (PEA), systems thinking and hypothesis testing. Although it is important to nuance this by reference to the *kinds* of intervention areas on which I4ID delivered results using this approach, it should help to inform current conversations in aid agencies about the analytical tools and monitoring, evaluation and learning (MEL) requirements for adaptive governance programmes.

Second, our study found that, despite aspiring to not invest in core funding or infrastructure for partners – keeping 'money off the table', in other words – most of I4ID's successes did involve the strategic investment of programme funds. Even in the small number of cases where brokering stakeholders was the principal engagement strategy, success was still accompanied by some form of capital investment or core funding.

1 Introduction

This report uses QCA to explore the most important causal ingredients of success in the I4ID Programme – an adaptive, issue-based governance initiative in Tanzania.

'Adaptive management' and 'politically smart programming' are increasingly popular ideas in development. They capture an ambition to programme in ways that are more flexible and experimental, and which respond to and capitalise on political dynamics and incentives. Over the last 10 to 15 years there has been a notable increase in the number of aid programmes that explicitly reference these terms, or what are said to be similar ideas, such as DDD and PDIA.

This trend reflects a recognition that, while some development priorities can be addressed by rolling out tried and tested solutions, many others involve addressing more complex and unpredictable issues. Consider, for example, reducing informal payments in primary health care provision (Kitson, 2019). There are human, cultural and political aspects to this, and it involves many social and economic factors. We can expect it to be a 'nested' problem, in that it most likely involves several inter-related and evolving issues – collective action, coordination, information asymmetry, and so on. Healthcare systems resist simple fixes; short-term improvements may be possible, but may not be sustainable over time without more fundamental systems change. Moreover, what works to reduce informal payments in primary healthcare in one context is unlikely to be directly transferable to another (Ramalingam et al., 2014). Finally, there is no obvious 'end point' where the challenge has been resolved.

Given that these kinds of challenges involve a range of moving parts, they are not susceptible to ready-made solutions or detailed planning, even if that planning draws on good technical knowledge. Rather than starting with a blueprint mapped out in advance, adaptive management calls for implementing teams to search for the right approach through a cycle of testing and learning, the results of which should then be fed back into programme implementation. That cycle might involve making a series of 'small bets' on a range of solutions and then pursuing only those which show the potential for good returns.

There is a growing consensus that interventions are more likely to make a positive difference in highly complex situations if they adopt these principles and methods (see, for example, Andrews et al., 2013; Faustino and Booth, 2014; Burns and Worsley, 2015; Green, 2016; Andrews et al., 2017; Kirsch et al., 2017). However, the development community is still at a relatively early stage in gathering and analysing evidence about how to do these things well, and the actual difference they can make in terms of delivering development impact. While there are interesting, well-written case studies in the literature, these do not yet constitute a strong evidence base that shows whether and how these efforts can be clearly linked to better results and more effective aid programming (Laws and Marquette, 2018; Dasandi et al., 2019). According to recent reviews of the literature on TWP and adaptive management, much of the evidence used so far to support these approaches is anecdotal, is not robust or comparative, and draws on a small number of self-selected, relatively well-known success stories (Faustino and Booth, 2014; Hudson and Marquette, 2015; Dasandi et al., 2016; 2019; Laws and Marquette, 2019). These studies tend to rely on interviews, document analysis or a form of action research, rather than methods more appropriate for establishing causal explanations. As a result, they tend to focus on the reform episode and/or the programming process, instead of providing evidence on how the intervention strategy contributed to particular outcomes (Dasandi et al., 2019).

The shallowness of the evidence base has two consequences. First, when it comes to programme design and decision-making, practitioners rely mainly on the combination of a few well-worn studies, folk knowledge and their wits; there is little in the way of systematic guidance. It is true that the literature highlights several recurring factors that are said to contribute to the success of more adaptive, politically informed programmes, for example that programme managers allowed local actors to take the lead; that there was a supportive environment in the donor agency; that there was a high level of political attention around the problem tackled by the initiative (Booth and Unsworth, 2014; Laws and Marquette, 2018). As we discuss in more detail in Chapter 2, these are among the causal factors that we have chosen to test in the case of I4ID.

Another issue is that, because of the paucity of well-documented cases, we know little about whether these combinations of ingredients travel well to other contexts (Hudson and Marquette, 2015: 74). For example, it may be the case that successfully incorporating politics and adaptive methods into programme design and implementation in the justice sector in a fragile and conflict-affected situation will involve a different set of factors in comparison to, say, a programme looking to reform urban transportation in a relatively stable context.

Next, proponents of adaptive programming lack a firm foundation from which to answer sceptics and critics within both their own agencies and potential partners, or among newcomers to the aid sector. With the UK following Australia and Canada to become the most recent major donor country to merge its development and diplomatic agencies, there is a need to articulate the value of the adaptive agenda in more robust ways not only to long-standing aid practitioners, but also to newcomers: diplomats, foreign policy experts and trade specialists.

As I4ID looks back at the results of its adaptive governance work in Tanzania, it is well placed to generate lessons that can help improve the depth of the evidence base in the ways described in this section. QCA was chosen for this study in light of the specific methodology and evidence gaps noted above, as it allows combinations of different causal factors to be tested against outcome results. This should generate insights into the ingredients of successful interventions in a more robust way than is afforded by anecdotal discussion of individual reform episodes or programming processes (Befani, 2016). As far as we are aware, this is the first time QCA has been used to investigate the success factors associated with adaptive and politically smart programming.

In Chapter 2, we provide background information on the programme. Chapter 3 outlines, in detail, our methodology and the analytical steps we took. In Chapter 4, we present our findings on the core ingredients of I4ID's successes. In the conclusion, we draw, from our results, recommendations for policy-makers, programme designers and programme commissioners in donor agencies interested in starting adaptive, politically smart governance programmes in Tanzania or countries like it.

A word of warning. First, this report demonstrates the ingredients for successful adaptive programming in Tanzania at a certain moment in time. We need to be a little cautious about extrapolating the findings to other settings, though we do provide observations in the conclusion on where we think the findings might plausibly transfer. Second, this report is not intended to demonstrate the value of adaptive programming vis-à-vis other approaches to development programming. It cannot say that, where I4ID succeeded, another, completely different type of programme would have failed, or that another type of programme would not have done better than I4ID. For those kinds of insights, a different type of study would be required.

2 Background to I4ID

I4ID was a five-year, FCDO- and Irish Aid-funded programme designed to promote inclusive development and strengthen democratic institutions in Tanzania. It was conceived during the presidency of Jakaya Kikwete, a period in which civil society and the media were becoming more active, government transparency was increasing and the political opposition was growing stronger. I4ID's original business case had three main pillars: 'deepening democracy', including support to parliament, elections and the media; 'coalitions and collective action', intended to 'facilitate the development of coalitions comprising government, the private sector and civil society which were to work together to overcome the obstacles to collective action on issues of shared and public interest'; and a 'research and learning' pillar, which would support the other two.[1]

The bid was won by a consortium led by multinational company Palladium, supported by SNV and BBC Media Action, all of which had a strong local presence in Tanzania, plus Global Partners Governance and ODI. The successful tender document envisaged a strong synergy among the different pillars, with coalitions and collective action at the centre.

The consortium was approved in August 2015, and began work in April the following year. At the beginning it struggled with a number of teething troubles, as well as a shock in October 2015, when the general election was won by the ruling party's John Pombe Magufuli, a relatively unknown political figure. It soon became clear that Magufuli's government would have much cooler relations with traditional development partners, that it would wage a 'war' against corruption, and that it was generally intolerant of media criticism, political opposition and dissent (Andreoni, 2018; Eriksen, 2018).

Through research and analysis, changes in team and consortium composition and encounters with the new regime, I4ID adapted to this constrained political space. Not without some debate and controversy (Green and Guijt, 2019), the programme's potential focus on political parties, MPs, elections and critical voice became muted; the focus on collective problem-solving, as well as a newfound interest in market and other complex systems, was deliberately amplified.

Other elements of the programme remained largely unchanged. I4ID staff were encouraged to identify potential issues via the media, politics or their own past experience and networks. By now under intense pressure to demonstrate results, they were asked to ideate novel areas where research and analysis, convening and brokering, prototyping and experimentation and media strategising might conceivably bring about inclusive development change, first on a small scale, and then for broader populations. Potential issues that passed internal I4ID and FCDO scrutiny were upgraded to 'active workstream' status. In this way, I4ID gradually developed a portfolio of issue-based workstreams.[2]

As Box 1 shows, there was little to ostensibly unite the different elements in this portfolio, which was sometimes confusing for observers, prompting accusations that it was just dabbling in a range of unconnected things. What, in practice, tied it together was I4ID's politically smart, adaptive way of working, which was somewhat distinctive and untested in the Tanzanian aid context. One of the aims of this report is to provide a proof of concept for this approach.

1 See https://devtracker.fcdo.gov.uk/projects/GB-1-204809/documents.

2 More detail can be found on I4ID's website: www.i4idtz.org

Box 1 I4ID workstreams

- Inclusive Education (IE), which worked in particular on getting the education system to respond to the needs of deaf children.
- Menstrual Health Management (MHM), which worked, among other things, to strengthen the market for more affordable menstrual health products.
- Regional Investment Facilitation (RIF), which sought to create an enabling environment for private investment, primarily in Iringa Region.
- Solid Waste Management (SWM), which promoted a variety of initiatives to improve waste collection in Dar es Salaam.
- Sunflower (SF), which explored ways of promoting sunflower production and processing.
- Urban Water (UW), which worked primarily with the water utility in Singida municipality to expand water access for peri-urban residents.
- Urban Women Vendors (UWV), which sought to improve the position of women vendors in Mwanza and Mbeya.
- Urban Spatial Development (USD), which brokered a new, participatory land planning process in the Dar es Salaam municipality of Kigamboni.

In addition to a lack of homogeneity across the portfolio, there was also heterogeneity within some workstreams. Each workstream typically included several, sometimes only loosely connected, initiatives or channels of activity, running either sequentially or simultaneously, and leading to defined outputs in the programme's Results Framework. To give an example, the SWM workstream included a pilot initiative to improve fee collection through geo-tagging; an experiment in using a Tractor-Trailer model to make trash collection more profitable; and an initiative to help drug-addicted wastepickers with rehabilitation.

In the programme's Results Framework, all of these outputs were expected to lead to 'outcomes'. I4ID had four outcomes against which it measured success:[3]

1. '**Policy improvements**' or, to be more specific, 'Recorded improvements in policy, policy implementation, and/or institutional arrangements with evidence that I4ID contributed to that change'.
2. '**Significant impact**' or 'Instances where improvements in policy, policy implementation, and/or institutional arrangements supported by I4ID will have a significant impact on inclusive development issues'.
3. '**Inclusive decision-making**' or 'Significant instances where democratic institutions (broadly interpreted) involved in the programme demonstrate behaviour that is consistent with a more inclusive decision-making, planning or policy process'.
4. '**Sustained inclusive decision-making**' or 'Significant instances where democratic institutions involved in the programme demonstrate sustained or repeated behaviour that is consistent with more inclusive decision-making'.

If one is familiar with the distinction in the development literature between inclusive outcomes and inclusive processes, 1 and 2 might be regarded as being about inclusive outcomes, 3 and 4 about inclusive processes (see, for example, Rocha Menocal, 2020).

Note that this study will only consider outcomes 1, 2 and 3. We believe it is too early to assess claims for 4 robustly, even though in some areas we see promising signs.

3 In the programme's Results Framework these were respectively called OC1.1, OC1.2, OC2.1 and OC2.2.

I4ID was wound up in 2020, a little ahead of schedule. This was not, apparently, because of underperformance – the programme had done well enough in several Annual Reviews and a Mid-Term Review[4] – but rather because of changing priorities on the part of the UK government and FCDO Country Office. Indeed, I4ID is worthy of our attention for the fact that, in some areas, it achieved real breakthroughs on previously neglected or intractable problems. To provide a couple of examples: despite in recent years an expensive donor-backed initiative to provide a policy and planning framework for IE in Tanzania (MOEST, 2017), this effort had yielded few tangible results on the ground. However, with only a small investment I4ID leveraged media exposure of a set of disastrous exam results at a school for the deaf into multi-stakeholder action which eventually led to the standardisation of a Tanzanian sign language (TSL), reform to the curriculum, examination system and teacher-training system and the potentially norm-changing use of signers for prime ministerial speeches and on TV current affairs programmes (I4ID, n.d.). To give another example, land use planning in the Dar es Salaam municipality of Kigamboni had been deadlocked for years following an exclusionary attempt to create a 'New City' in the area, which had eventually fallen foul of political protest (Lindell et al., 2016). I4ID, through a combination of politically smart, convening and brokering activity and judicious technical inputs, was able to bring previously hostile actors together in a new, more inclusive land use planning process, which was eventually incorporated into the Dar es Salaam Masterplan and adopted by the World Bank as a model for other cities.[5]

A fair amount has already been written about I4ID,[6] yet to date there have been no systematic and rigorous attempts to capture the ingredients behind its success. This report attempts to rectify that.

4 These reviews can be found at https://devtracker.fcdo.gov.uk/projects/GB-1-204809/documents.

5 www.I4IDtz.org/wwwo-urbanplanning.

6 In addition to a political settlements analysis of Tanzania (Kelsall, 2018), ODI produced several short briefing notes looking at I4ID's approach to partnerships, value for money and gender mainstreaming (Laws, 2020a; 2020b; 2020c). Green and Guijt (2019) produced a longer standalone case study on the programme as part of the Action for Empowerment and Accountability research programme.

3 Methodology

This report uses QCA to explore the causal ingredients behind I4ID's success in Tanzania. QCA uses set-theoretic logic to synthesise data and identify the patterns behind successful outcomes, including whether particular ingredients or conditions, or combinations of ingredients or conditions, are necessary or sufficient, or merely important, for producing the desired outcomes.

3.1 The rigour of QCA

QCA provides a systematic way of identifying the ingredients of programme success, providing a more rigorous approach than previous analyses of adaptive programmes in at least two ways.

First, QCA provides a way of reducing complexity in a dataset and allowing systematic, as opposed to intuitive, comparisons to be made across a large number of cases (Befani, 2016).

Second, QCA is more transparent. Most studies of adaptive programmes have used an inductive or abductive approach[7] to synthesising data, picking out what seem to be underlying patterns. While such analyses can be very insightful, there is always the danger of confirmation bias, as researchers cherry-pick from the data patterns that support their pre-existing assumptions and beliefs. QCA does not completely remove the problem of confirmation bias, since researchers still have to pick and choose what to highlight, often from a vast number of results. However, it is more transparent than other methods, since the dataset on which the analysis is conducted, together with the various models that have been tested, are typically made available. Other researchers can then try and replicate the findings, run other models or choose to highlight other findings.[8]

We have chosen to place all our technical results in an Annex, and have made it easy for readers to see where, from among this plenitude of results, the findings we have chosen to highlight have been drawn.

3.2 Basic approach

QCA is a highly technical exercise, but the ideas behind it are quite intuitive. Imagine you are planning to open a restaurant in Tanzania and you want to know the ingredients for making a successful pilau, a popular local dish. You are familiar enough with the local cuisine to know that there is no single right way of making a pilau, and that a variety of ingredients can be used, with some more common than others.[9] To maximise the chances of your restaurant being a success, you decide to ask 20 chefs to each cook their own pilau, with the results judged by a panel of would-be customers. You then look for the combination or combinations of ingredients that lie behind the successful dishes. You also have a sense that the ingredients can be divided into several broad categories: rice; spices; herbs; vegetables; meat; and also that several cooking skills might be important, such as the ability to plan, time, judge proportions, taste or follow a recipe, or prior experience of cooking pilau.

7 See https://en.wikipedia.org/wiki/Abductive_reasoning.

8 It is also worth noting that Barbara Befani, who ran the technical analyses, had no close familiarity with the programme and no particular agenda with respect to adaptive programming. The patterns that emerged from the data analysis stage were about as free from bias as it is possible to be.

9 A brief survey of pilau recipes online reveals the following ingredients: rice, cinnamon, cumin, star anise, bay leaf, curry powder, cardamom, black pepper, garlic, ginger, onions, bell pepper, Jalapeno pepper, chilli flakes, coconut milk, garam masala, bay leaves, tomatoes, potatoes, lamb, chicken, saffron, turmeric, cloves, coriander, beef.

A QCA analysis codes these different ingredients as either present (1) or absent (0). So, we might say that, if a dish had four or more spices, we would code spices '1' (ingredient is present), and if it had three or less, we would code it '0' (ingredient is absent). And if a dish were deemed successful, we would code it 1, and if unsuccessful, 0.

QCA uses three different techniques to analyse the relationship between these ingredients and success or failure (Befani, 2016). The first is superset analysis. This looks across the set of successful cases and asks whether there are any ingredients that, on their own, appear in every successful combination. If there are, these ingredients are judged 'necessary' for success.

The second is subset analysis. This looks at the individual ingredients first, then successively at combinations of two, three and so on, to find whether they are associated with success, with failure, or both. If they are only associated with success, they are judged 'sufficient' for success, because whenever that combination is present, success is also present.

The third QCA technique we use is Boolean minimisation. What Boolean minimisation does is to successively eliminate ingredients that are not equally present in otherwise identical combinations that present the same outcome, deeming them not to be causally crucial because the outcome and everything else is the same in the two combinations that are being compared. So let's say, simplifying somewhat, that some of our successful pilau had potatoes, some of our successful pilau did not, some of our unsuccessful pilau had them, and some of our unsuccessful pilau did not. We infer that the presence or absence of potatoes is not of critical importance to the success of a pilau: we eliminate potatoes from our analysis, and narrow our focus to the combination of ingredients that remains consistently associated with the same outcome.

We will be doing something similar in this report, albeit with a much more complex phenomenon. The idea is to show policy-makers interested in starting adaptive programmes in Tanzania or similar countries the ingredients most meaningfully associated with a track record of success: the core ingredients, in other words.

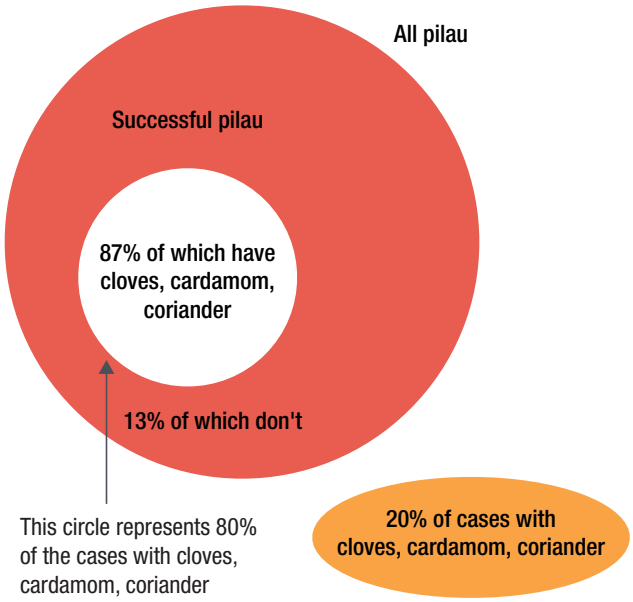
A word of warning: the dataset we will be looking at contains a very large number of ingredients across a very large number of outputs with three different kinds of objective (to return to the pilau analogy, we have dozens of ingredients, more than 60 pilau to evaluate, according to three criteria, for example how flavourful, nutritious and filling they are). Because of the size and complexity of the dataset, it is impossible to analyse it all at once. Rather, we go step by step, subdividing it according to our main categories of ingredient, and then testing different combinations of ingredients, sequentially, within those categories. In QCA terminology these combinations are called models, and they might be thought of as hypotheses about what is going to be successful. To follow our cooking analogy, we might say, 'I'm guessing that cloves and cardamom and coriander is a good combination, so let's analyse the results for all those dishes that combine them and see whether they succeed in terms of being flavourful, nutritious and filling'. And we might find that most of the pilau that have this combination are indeed flavourful, but it is harder to see a relationship to being nutritious and filling. We then ask what percentage of successful dishes is covered by that combination (coverage), and also what percentage of that combination (consistency) results in success.

In Figure 2, we find that cloves, cardamom and coriander have an 87% successful coverage with 80% consistency when it comes to flavour.

We then try another combination, for example cloves, cardamom and star anise. Does this have better coverage, or better consistency?

In this way we build up an iterative picture of the ingredients of success. Note that some models will have better coverage and consistency than others, which means that we are likely to rely on them more when drawing conclusions. But this does not mean that other models are wrong; indeed, they might reveal noteworthy combinations that 'more successful' models do not, and which are consequently worth referring to. In the example above, it might be that cloves, cardamom and star anise turns out to only cover a small percentage of pilau, but those it does cover are *all* successful, making it a combination worthy of our attention.

Figure 2 Stylised example of coverage and consistency in QCA results



Source: Authors' elaboration based on the survey results.

Having explained the basic idea behind QCA, we now proceed to discuss the different stages in the analytical process.

3.3 Choosing the units of analysis

At the outset we had a choice to make about what our basic units of analysis or cases would be. As mentioned above, I4ID organised its work under several different workstreams. Each of these workstreams contained several different, not always closely related, initiatives or activities intending to lead to a conclusion, with varying degrees of success. As mentioned, the SWM workstream included a pilot initiative to improve fee collection by means of geo-tagging, and another initiative to make refuse collection more profitable by introducing a tractor and trailer approach. Both were concerned with SWM, but in different ways. Moreover, once the geo-tagging initiative got beyond the pilot stage, scale-up of the model demanded a new set of activities. This kind of variation made an overall evaluation of workstream success a little artificial. We solved this problem by turning to the programme's Results Framework. Here, individual initiatives were expected to lead to designated outputs, such as 'successful completion of fee collection pilot', which were then expected to lead to or constitute one or more of the programme's outcomes (that is, 'Policy Improvement', 'Significant Impact' or 'Inclusive Decision-making').

In total, we analysed 65 different initiatives or outputs, which we use as our basic units of analysis, or cases, listed in the Annex. From here on, we use the terms 'output' and 'case' interchangeably. We assess these cases' success or failure by reference to their performance with respect to four outcomes discussed below.[10]

One of our reviewers noted that, because I4ID's workstreams were so different, we are not really comparing like with like, and, indeed, that it was much simpler to achieve success in some workstreams, where actors were broadly aligned around a common goal, than others, where there were serious conflicts of interest. To quote, one was like 'boiling an egg', the other 'cooking Beef Wellington'. While there is some truth in this observation, we believe that, once workstreams are broken down to the output level, which tended to capture incremental progress on an issue, the differences are less stark, and the cases are therefore more comparable.

3.4 Choosing the ingredients

The next step in our process was to compile a list of potential ingredients that might plausibly lead to success (or, in a few instances, be expected to bring failure). Here, we relied on the background literature on adaptive programming, which identifies a variety of ingredients, such as the need for local leadership, the desirability of a 'money off the table' approach, the importance of political analysis and the need to be iterative. We combined this with our own observations of the programme, which highlighted, for example, the use of a

10 For a comprehensive overview of the extent to which our cases were successful, readers can examine the Outcomes tab on the spreadsheet that accompanies this report, where the cases appear in column A and the outcome results in columns B–E. Alternatively, readers can study the Venn diagrams in the Annex.

'systems approach' in some workstreams and the fact that at times the programme, against the literature's advice, did provide capital investment. We then went through several iterations of brainstorming with programme staff, asking them to pick apart our list of ingredients or add others. Finally, we asked our QCA expert, Barbara Befani, to refine the definition of ingredients as she saw fit, which led to the subdivision of some and the amalgamation of others.

The final list contained 43 different ingredients (Table 1). It represents a compromise between being 'mutually exclusive and collectively exhaustive' in a logical and empirical sense, and listening to what the programme team in Tanzania thought was important.

3.5 Coding the ingredients

Coding the ingredients was the next and most laborious step. For each case we asked whether the different ingredients were present, using a combination of internal document review and interviews with workstream coordinators or workstream leaders. At the outset information was captured qualitatively, with examples and justifications recorded where appropriate. Next, the information was coded in binary fashion. This was arguably the trickiest part of the exercise, involving judgement calls on the part of the researchers, sometimes in dialogue with Programme staff, about whether to code an ingredient as present or absent. Although some of our ingredients were like spices, and could be easily coded as 1 or 0, many were more like 'skill', and required a more subjective assessment. Where we were not sure, either because there was not enough evidence on which to make a clear assessment or the evidence was indeterminate, we coded the case with an asterisk[11] to connote our uncertainty, but also indicate in which way we thought the evidence leant. This affected a significant proportion of cases, perhaps 30%.

Note that where two or more ingredients[12] for a particular output within an overarching condition were coded as uncertain, we did not include that output in the models for that condition. To provide an example, under the Politics overarching condition, case number 8, the scale-up of a mobile transfer station initiative in the SWM workstream, was coded as 'uncertain' for ingredients 1.3 and 1.4. Because of this uncertainty, we excluded this case from the analysis of the political ingredients of success, relying instead on other outputs where the data was less ambiguous. For that reason, this 'case' does not appear in the Venn diagram results displayed in the Annex. As a rule, where cases are missing from the Venn diagram results, this is why.[13]

3.6 Coding the outcomes

We coded the outcomes in a similar fashion. Readers will recall that I4ID had four outcomes in its Results Framework, but that we will only discuss the first three, namely 'Policy Improvements', 'Significant Impact' and 'Inclusive Decision-making'. To assess whether a case was successful against these criteria, we relied on claims the programme itself made in its Results

11 With respect to our case coding, 1* and 0* indicates that, while we are not completely confident, due to a lack of evidence or lack of clear evidence, we feel the balance of evidence points respectively to a positive or negative coding. Eagle-eyed readers may also notice in our spreadsheet a handful of '0.5' codings, where we were genuinely undecided. Out of an abundance of caution these were also excluded from the analysis. We also coded some cases N/A on particular conditions. This indicates that, while a particular ingredient was not present in that case, its absence was (a) not the result of an explicit decision on the part of the programme team; and (b) irrelevant to the success, or otherwise, of the particular intervention.

12 On the Capacity tab, it was one or more condition.

13 Given the nature of the data, some QCA experts may ask why we did not opt for a 'fuzzy set' analysis. The simple answer is that fuzzy sets, when done properly, are more difficult to calibrate and make sense of than crisp sets; moreover, the returns are debatable, since for the Boolean minimisation one cannot use fuzzy sets, and must 'crisp' the dataset before synthesis. We feel our strategy of removing the uncertain cases and only analysing the values that could confidently be crisp is the more robust option, both in terms of transparency and control over calibration and interpretation of findings, and efficiency.

Table 1 QCA conditions and ingredients

Overarching condition	Specific ingredients
1. Politics. The initiative paid attention to its authorising environment and/or employed politically smart programming principles	1.1 A reasonably in-depth study of some sort was commissioned[i] 1.2 An ex ante in-depth political economy study was conducted 1.3 The initiative responded to a problem already receiving a high level of political attention 1.4 The initiative responded to a problem already receiving a high level of media attention 1.5 The team responded to or sought out political stakeholders with a potentially positive interest in the initiative 1.6 There was ongoing political engagement and light-touch everyday political/political economy analysis 1.7 As the programme evolved, more in-depth political economy analysis studies were commissioned 1.8 Workstream coordinators or other team members were given space, time and resources to follow their 'political antennae'[ii] 1.9 The authorising environment for donor support and funding was maintained.
2. Design. The initiative employed PDIA/human-centred/systems design principles	2.1 The team engaged in rigorous systems research, mapping or modelling around the initiative 2.2 The team took a less formal systems approach – recognising the complex, multidimensional nature of most problems and searching, iteratively, for solutions, but not using specific systems research, mapping or modelling tools/methods 2.3 The team approached the problem as a discrete, one-dimensional issue for which the solutions could be pre-determined, pre-planned and implemented more-or-less as intended (i.e. there was no systems thinking in either a formal or organic sense) 2.4 The team took a 'human-centred approach' – i.e. involving participatory action research, deep immersion in context, community brainstorming, usability scales, sustained community feedback to inform decisions on scale, etc. 2.5 Solutions were ideated through processes of positive deviance or latent practice 2.6 There was a deliberate attempt to adapt external best practice or technology to the local context
3. Actors. The initiative leveraged or paid attention to local leadership/ownership/ acceptance	3.1 The initiative received 'buy-in' from senior politicians 3.2 The initiative received 'buy-in' from lower-level politicians 3.3 The initiative received 'buy-in' from implementing agencies (ministries, executive agencies, local government) 3.4 The initiative focused on problems that mattered to local non- state actors, for example civil society, citizen groups or the private sector 3.5 The workstream faced opposition from well-coordinated and/or politically influential actors or groups
4. HR. The initiative employed appropriate staff or engaged a wider team with what might reasonably have been expected to be the requisite abilities, political networks or technical skills normally associated with this problem area	4.1 Team or coalition members have strong local networks 4.2 Leading team or coalition members had been resident in Tanzania for many years 4.3 Team or coalition members have held senior positions in a national organisation 4.4 Team or coalition members are widely known in this geographic, issue or policy area 4.5 Team or coalition members are connected to government and/or civil society networks 4.6 Team or coalition members are strongly identified with an opposition party 4.7 Coalition members have a successful track record in political analysis, Gender and Social Inclusion (GESI) analysis, market systems analysis or other relevant fields of technical knowledge 4.8 Team or coalition members have a successful track record in using adaptive approaches
5. Learning. The initiative employed principles of iterative adaptation	5.1 The initiative deliberately engaged in at least a loose kind of hypothesis formation upfront 5.2 The initiative identified multiple rival hypotheses upfront and tested them in parallel 5.3 The initiative clearly and explicitly identified the conditions for testing hypotheses, with clear success/failure criteria 5.4 The initiative subjected its hypotheses and prototypes to a rigorous and structured process of implementation and testing 5.5 The team went beyond purposive muddling, to systematically analyse and understand successes, setbacks and failures against initial hypotheses, enabling modification and adaptation 5.6 A number of experimental, iterative steps were progressively employed to enable real solutions to emerge 5.7 There was a clear plan for how initial success would lead to scale-up
6. Funding. The programme made funding decisions that were consistent with adaptive management conventional wisdom	6.1 The initiative deployed programme funds in a smart or strategic way, i.e. as a catalyst for significant change, or to provide proof of concept for a pilot or solution that could be scaled up without further I4ID investment 6.2 The initiative took a 'money off the table' approach, i.e. funds were used principally for brokering and convening, instead of funding pilots, infrastructure, technical assistance, or to provide core funding 6.3 The Programme agreed to provide funds for pilots, infrastructure, technical assistance, or core funding 6.4 The resources made available in the programme budget were regarded as sufficient by the workstream co-ordinator
7. Capacity and innovation. The initiative sought to innovate or to improve existing solutions	7.1 The initiative deliberately tried to introduce and/or build capacity for an innovative technology 7.2 The initiative deliberately tried to introduce and/or build capacity for an innovative behaviour 7.3 The initiative sought to 'fine-tune' an existing system, model or technology, rather than look for an innovative solution, model or approach

i With the benefit of hindsight, we think this ingredient might be more suitably placed in the Design category. However, we do not believe it has had a large impact on the overall results.

ii This phrase was used by Green and Guijt (2019) in their case study on I4ID. It refers to the kind of heightened political intuition or awareness, possessed by some staff members, who have typically been immersed in a particular political environment for a significant period of time, and who understand intimately both the formal and informal 'rules of the game'.

Framework, nearly all of which had been checked by external evaluators.[14]

It is important to note that, being an adaptive programme, I4ID did not have many 'failures'. In most cases, initiatives deemed to be at risk of failing were either dropped or adapted. Consequently we set a high bar for success, coding some of these successes with a variety of asterisks out of an abundance of caution.[15] Although there is a possibility that some I4ID Programme staff might take exception to these codings, perhaps believing them to impugn their efforts, we think this is a justifiable price to pay for basing our conclusions on results in which we are truly confident.

We then ran a variety of tests on our data, in which we included these starred results to different degrees. In effect, this led to a recalibration of our three outcomes into four:

1. Policy Improvements (from which we excluded all the 'uncertain' codings)
2a. Significant Impact – low bar (from which we again excluded all the uncertain codings)
2b. Significant Impact – high bar (in which we excluded the 1* codings but counted the 0* and 0** codings as 'failures')
3. Inclusive Decision-making (in which we also excluded the 1* codings but counted the 0* and 0** codings as 'failures').

A virtue of doing things this way is that the number of 'failed' cases increases steadily across the four outcomes, meaning that our hypotheses about the ingredients for success were subject to a series of increasingly strenuous tests.

3.7 Highlighting findings

Readers will see from the Annex that we ran a large number of tests on the data and the results can seem overwhelming. How then did we choose what to highlight? In brief, we followed three rules of thumb:

1. Superset results that demonstrated that certain individual ingredients were necessary for success, though with a caveat about triviality (see below).
2. Boolean results that demonstrated that certain combinations of ingredients were important for success. Usually, this would mean that 90% or more of successful cases on the more challenging Outcome 2B or 80% or more on Outcome 3 would be covered by this combination.
3. Boolean results that demonstrated that certain combinations of ingredients were consistently successful across all outcomes.

Proceeding along these lines, we would still sometimes arrive at an almost indigestible set of combinations. To summarise further, then, we try to generalise across the findings, boiling them down to their essentials, trying to extract what is really significant for would-be programmers.

To return to our pilau analogy, and only looking at spices, we might find that cloves were present in all successful dishes, but beyond that, successful combinations could include cloves with cardamom, cumin and coriander; or cloves with coriander and black pepper but not cardamom and cumin; or cloves with cardamom

14 An external Results and Challenge team had verified the achievement of each output for the Annual Review up to 2019, and further outputs were verified as part of the Mid-Term Review.

15 For our outcome coding, we tried to reflect subtle but important distinctions in how the programme could be said to have succeeded or failed with respect to particular results, and the confidence with which we can make these judgements on the balance of evidence available at the time. 0 indicates that the programme tried to contribute to this outcome through this intervention but failed. 0* indicates that the programme tried to contribute to this outcome through this intervention, but there is insufficient evidence to support claims for success. 0** indicates that the programme tried or considered trying to contribute to this outcome through this intervention, but then dropped it because it suspected it might fail. 1* indicates there is some evidence the programme was successful in contributing to this outcome with this intervention, but we recognise there may be reasonable disagreement about whether a sufficiently high bar was met. 1 indicates that the programme tried to contribute to this outcome through this intervention and succeeded. N/A indicates success on this outcome was not an intended part of the theory of change for this intervention.

and star anise, but not cumin and coriander or black pepper. Rather than advising our chef to follow these combinations precisely, we instead try to generalise, saying something like: 'Cook with cloves, and – provided that none of these combinations is associated with failure – at least one other of the aforementioned spices'.

Now to the triviality matter. Some of I4ID's ingredients were present in all cases. They therefore appear, in some combination with other ingredients, in all successful and unsuccessful outcomes, and thus appear to be necessary for both success and failure, and sufficient for neither. Logically, the necessity for success is trivial, since the ingredient is also necessary for failure. But that does not mean that it is empirically trivial, since we have no empirical evidence about what happens when this ingredient is absent. Suppose, in the case of our pilau, all the dishes, successful or not, included rice. One could, in one's advice to a would-be restaurateur, say, 'Don't bother with rice, it's trivial'. But anyone who has ever eaten a pilau would know this to be a mistake. Thus, on grounds of risk avoidance, we advise practitioners to begin by including those ingredients that were always present in I4ID and which were thus a constant in its successful initiatives. At some point they may wish to experiment with leaving these ingredients out, but that would not be a wise way to start.

3.8 Transferrable lessons?

To what extent might these ingredients for success be transferrable beyond Tanzania? Although QCA of this nature imparts a strong measure of 'internal validity' to the findings, it provides no guarantee of external validity. We cannot generalise beyond the Tanzanian context with any certainty. However, it seems reasonable to hypothesise that a programme with these ingredients should deliver successful results across a range of low- and middle-income settings with a moderate level of state capacity, civil society activity and private sector development and, perhaps most interestingly, within constrained political space.

4 Main findings

4.1 Meta-analysis

Our first step was to conduct a meta-analysis of the broad categories of ingredients that went into the I4ID programme.

Here, an ingredient was coded 'present' if a majority of its sub-conditions (suitably calibrated) was also present. For example, under the Politics category, there were nine ingredients, so if five were present, we coded Politics as 'present', and if only four, we coded it 'absent'.

Proceeding in this way, we found that 'Funding' was necessary for success on all four outcomes, and that 'Politics' was necessary for success on Outcome 3, 'Inclusive Decision-making'.[16] This might be explained by reference to the fact that success always requires a balance of appropriate funding strategies, and that Outcome 3 is a more inherently political goal than the other outcomes.

That said, we think it is important not to read too much into this meta-analysis. Because of the way we have constructed our ingredients, 'absence' of Politics on the first three Outcomes does not imply that political ingredients can be dispensed with completely. The same applies to the other, 'non-necessary' ingredients.

Boolean analysis reveals a number of combinations – some of which involve additional ingredients – that are consistently successful across all outcomes.[17] But again, we do not want to over-emphasise their significance. More light is provided, we believe, by the within-category analyses that follow this section.

As with our pilau analogy, one wants to know, not just that spices are necessary, but which ones.

4.2 Politics

4.2.1 Summary

I4ID was designed to think and work politically to deliver inclusive development and strengthen democracy via institutional change. In consultation with the wider literature and team members, we identified eight different potential political ingredients that a programme like I4ID might have used in this endeavour. Having applied our rules of the thumb to the findings that were generated by running various tests on our dataset, there are four models which we believe are worth highlighting, which we describe and illustrate below. Synthesising across these models, we recommend that future programmers in places like Tanzania should:

- address problems already receiving political attention
- respond to or seek out political stakeholders or champions
- review progress using light-touch political analysis tools, perhaps with more in-depth analysis at critical junctures
- maintain the political support or 'authorising environment' of the funder.

4.2.2 Models, illustrations and discussion

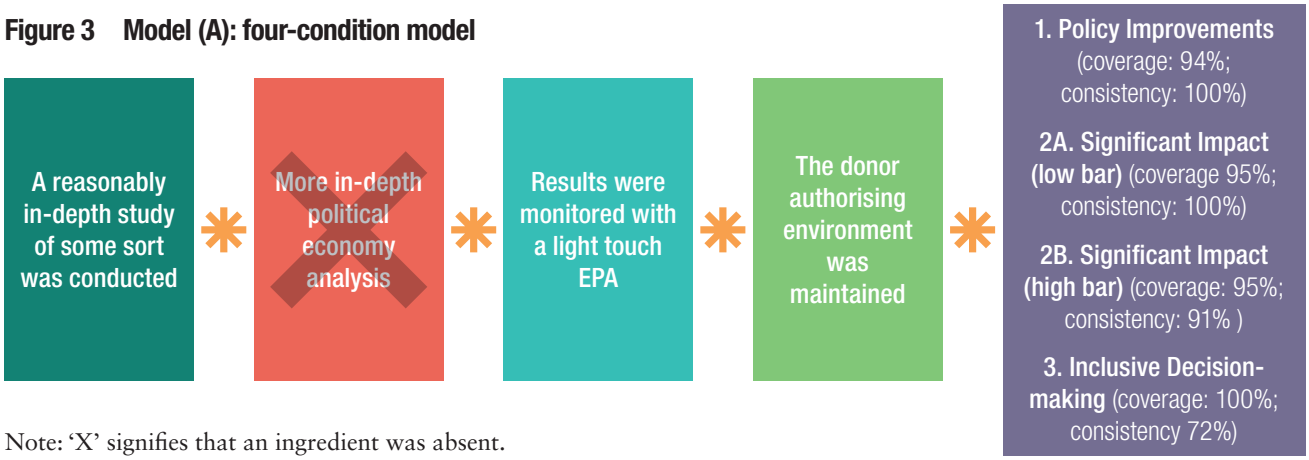
Model (A)

We began by testing a nine-condition model, then narrowed this down to a six-condition and then a four-condition model. The four-condition model confirmed that, in a large majority of successful outcomes, studies were commissioned (1.0), there was no upfront PEA (1.1), results

16 See Annex, p. 10, para 2.0.

17 See Annex, p. 10, para 3.0.

Figure 3 Model (A): four-condition model



Note: 'X' signifies that an ingredient was absent.

were monitored with a light touch EPA (1.5), and the donor authorising environment was maintained (1.8).[18] To provide an illustration, one of the programme's workstream targets was that a new Regional Investment Facilitation Office (RIFO) it had helped create in Iringa would facilitate at least one new investor in the region. This was deemed to have been achieved when, in February 2020, HongWei International, a Chinese company that manufactures plywood in Iringa, opened a metal fabrication workshop, with the expressly acknowledged support of the RIFO. With other investments in the pipeline that might plausibly bring 250 good jobs to the region, the programme claimed success on its second 'Significant Impact' outcome indicator.

To get there, the programme had, among other things, conducted a study of investment in the region and the main bottlenecks facing investors (though this was not explicitly a PEA study) (1.0); it had held regular, ongoing discussions in weekly team meetings, and informally, about its unfolding progress through a political lens, discussing in particular the political motivations of the Regional Commissioner and other stakeholders in the region; and the workstream had used political analysis tools, including stakeholder analysis, power cube analysis and change-space analysis, in a light-touch way on more than one occasion (1.5). The programme had also been able to maintain the 'authorisation' of donors, despite concern at times that progress was not swift enough or the investments not inclusive enough, partly by taking donors on site visits (1.8).[19]

However, this combination of ingredients was not always associated with success, showing inconsistent results with respect to Outcomes 2B and 3 (see Figure 3), and other combinations could also be successful. Consequently, we tested other models with other combinations of ingredients, to see if they could also contribute to our understanding.

Model (B)

Emerging out of these analyses was a second model we believe worth highlighting (see Figure 4): studies were conducted (1.0); the problem was already receiving political attention (1.2); and political stakeholders were deliberately sought out or responded to (1.4), but *without* the help of a more in-depth PEA (1.6). A significant majority of successful cases included this combination – albeit with only marginally higher consistency on Outcome 3 than Model A.[20]

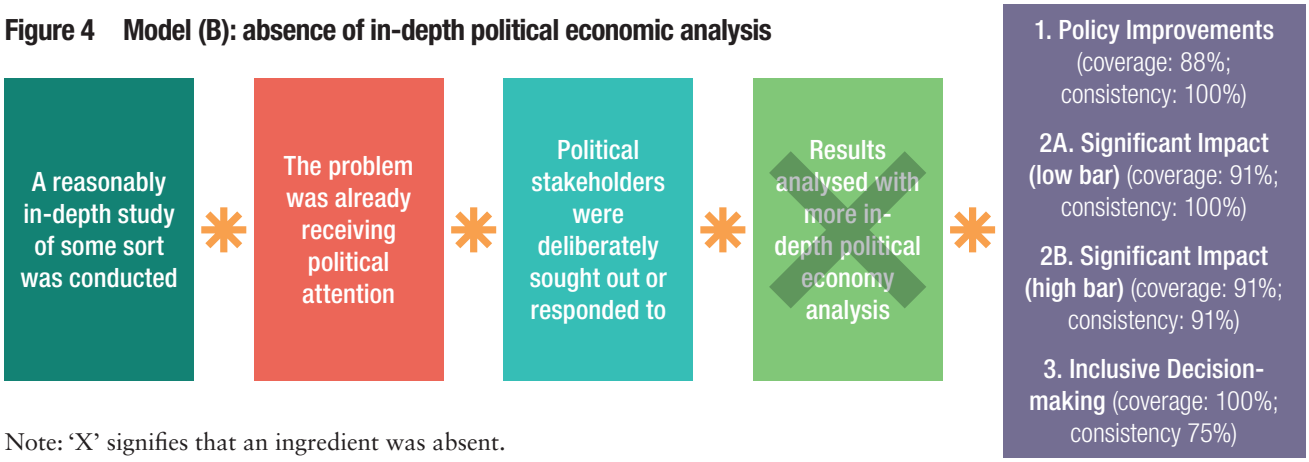
In the case of the RIFO workstream, for example, the programme was responding to a

18 Annex, p. 19, para 9.1.

19 Annex, p. 19–20, Figs. 21 and 22, RIF OP1 (1) in Box 1011.

20 Annex, p. 25, paras 12.0–12.1.

Figure 4 Model (B): absence of in-depth political economic analysis



Note: 'X' signifies that an ingredient was absent.

situation where the president was pushing an industrialisation drive and had explicitly directed Regional Commissioners to stimulate industrial development in their regions. Through contacts at the Ministry of Local Government, the programme had been directed to Iringa Region, which was thought to be fertile ground for its support. Because of this, the programme was very much going with the grain of national and local politics.

It is instructive in this respect to examine an obverse case in which the first three of these ingredients was absent. As an example, one of I4ID's least successful initiatives was an effort to increase advocacy and improve working conditions for waste-pickers at Dar es Salaam's Pugu dumpsite.[21] The initiative was abandoned after it became clear that ruling party stakeholders on the municipal council regarded it as a potentially dangerous issue that might play into the hands of the political opposition. Indeed, the absence of these ingredients, though rare, was consistently associated with failure on Outcome 3.[22] Another rare but interesting case, illustrating another variation of these four ingredients, was a pilot scheme to test a new method for collecting waste at a very local level in Dar es Salaam.[23] Although the programme had sought support for the pilot from street-level political leaders and members of the relevant municipal technical department (1.4), it was abandoned when higher-level political forces changed the regulations around revenue collection and retention, rendering the scheme unviable. A plausible inference is that the choice of political champions is important: it helps to have the most powerful stakeholders onside.

Model (C)

Despite the explanatory strength of Models A and B, they do not reveal any *consistently* positive pathways to Outcome 3, Inclusive Decision-making. The one consistently successful combination of ingredients leading to this outcome emerged from the following five-condition model (Figure 5). This involved the presence of commissioned studies (1.0), political attention (1.2) and engagement with political stakeholders (1.4), but *not* media attention (1.3) or robust political economy studies (1.6).[24] This combination was found in a range of outputs in the UW and RIF workstreams.

It is interesting that the *absence* of in-depth political economy studies (1.1 and 1.6) tended to be *positively* associated with success.

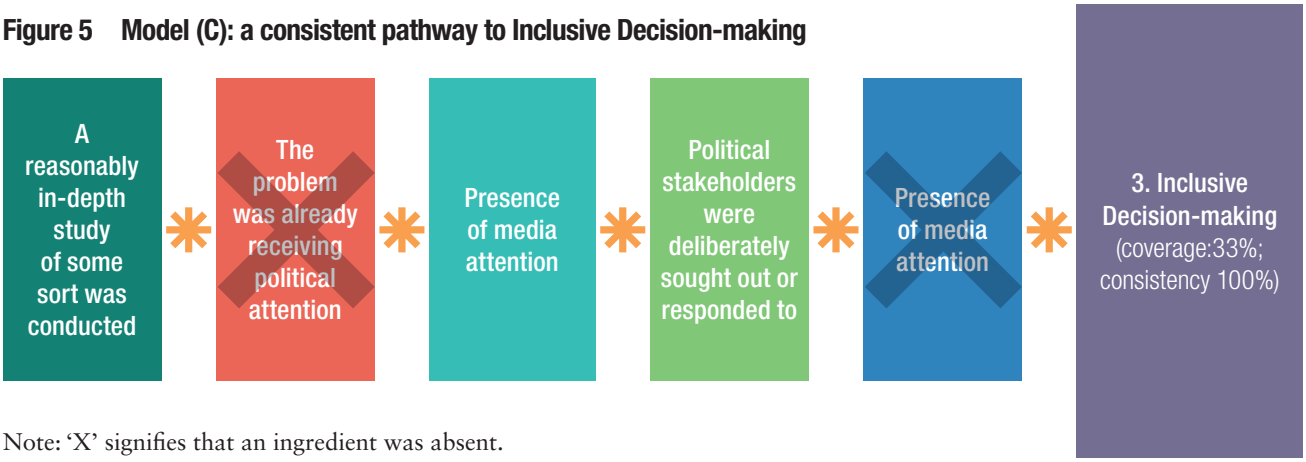
21 10. SWM8

22 Annex, pp. 26–28, Figure 35, top-left rectangle coded 0000.

23 No. 11, SWM 10. See Annex, Figure 34, left-sided rectangle coded 0010.

24 Annex, p. 23, paras 11.1–11.2; p. 25, Figure 31, green box in the bottom right, labelled 11010.

Figure 5 Model (C): a consistent pathway to Inclusive Decision-making



Note: 'X' signifies that an ingredient was absent.

While acknowledging that we have a bias here, we believe this finding should be caveated.[25] While it is true that upfront PEA of workstream or workstream outputs was not conducted, a detailed analysis of the political economy of the country context *was* undertaken in the programme's inception phase, followed by a further one midway through Magufuli's first term. These helped orient the programme in a general way, although their advice was not always followed, with programme staff, especially the team leader, often preferring to follow their political instincts, rather than analytical guidelines.

Only in one workstream was an in-depth political economy study conducted as the workstream unfolded. This was USD, which commissioned a locally led PEA study when it began to encounter political opposition. The workstream went on to enjoy success across a large number of outputs, and there are reasonable grounds for thinking that the PEA study was a contributing factor insofar as it helped persuade council stakeholders to include in the planning process a potentially disruptive citizens' group they had previously been determined to sideline. Further, no cases where an in-depth PEA was conducted were unsuccessful.

In general, however, the team leader tended to treat greater-depth PEA as a luxury ingredient – a spice, perhaps like saffron – that the programme could not really afford. Because it was employed so sparingly, it is difficult to know whether or not it would have improved results or ended up being good value for money, or indeed whether it would be necessary in other cases, like USD characterised by deeper collective action problems or conflicts of interest. We are satisfied that in most cases, however, I4ID demonstrated that it could be successful without it.

Media attention was another ingredient that did not emerge as particularly significant. In most cases, problems surfaced via the team members' personal contacts or networks, not media publicity.[26] Again, however, we wish to caveat these findings. As one of our reviewers pointed out, we would expect media work (and, we would add, in-depth PEA) to be important for the broader replication and adoption of policy change. This goal is captured in the fourth outcome in the I4ID Results Framework, on Inclusive Decision-making. However, readers will recall that we opted not to include this outcome in our analysis as, for the most part it is too early to assess whether outputs have contributed to it.

25 ODI specialises in PEA and one of the authors was PEA Advisor to the programme.

26 Media attention was created and leveraged in some successful workstreams, mainly with the help of BBC Media Action, a consortium partner.

Finally, it remains true that workstream coordinators were allowed to follow their political antennae on most occasions. However, according to our rules of thumb, the QCA analysis reveals that other factors were more significant. Reading across these models, four ingredients are consistently associated with outcome success, and we have highlighted these at the start of this section.

4.3 Design

4.3.1 Summary

In diagnosing problems and designing interventions, I4ID had ambitions to take a 'systems approach'. In contrast to simpler forms of problem analysis which focus on one dimension – say, low state capacity – a systems perspective regards some problems as the product of a more complex, non-linear and unpredictable interaction between different social, economic and political dimensions (Ramalingam et al., 2008).

We identified six potential ingredients under this condition, which reflect the extent to which interventions might follow a more or less rigorous systems approach (or adopt a simpler kind of analysis), and whether specific complementary tools and methods for problem definition and intervention design are used. We tested various models on the dataset, and applied our rules of thumb to identify particularly successful ingredients and combinations. Generalising across these findings, there are two core ingredients for success we would recommend for future programme designers:

- using a light-touch systems approach (which is particularly effective when combined with positive deviance and/or leveraging external best practice), and treating more rigorous research as an optional extra
- using human-centred design, particularly when trying to bring about more inclusive decision-making (Outcome 3).

4.3.2 Models, illustrations and discussion

Only two workstreams conducted rigorous systems research or mapping at the outset. MHM conducted a rapid market systems analysis to understand supply chain issues and consumer needs, while the USD workstream brought together technical specialists, local and central government representatives, citizens' groups and vulnerable communities in Kigamboni to map land use. Both these workstreams had mainly successful outputs.

However, when we look at the dataset as a whole we find that a lighter-touch approach (2.2) featured in a significant majority of cases that successfully delivered outcomes. We interpret this to mean that rigorous systems research should be regarded as a potentially useful optional extra in a core ingredient combination, but not central to success (somewhat similar to in-depth PEA).[27]

When we focus on Outcome 3, our analysis indicates that either leveraging positive deviance/latent practice (2.5) or adapting external best practice (2.6) was required.[28] As an example of the former, during the design phase of the UWV workstream, I4ID partnered with a local women's rights organisation, Kivulini, to conduct an initial situational analysis and stakeholder mapping in Mwanza. Kivulini also advised I4ID that it would be important to obtain buy-in from local stakeholders early on, and coordinated a series of consultative workshops with I4ID's support. These consultations brought together representatives from the local government association and a 'machinga' (urban vendor) association, SHIUMA, to agree to work together on issues affecting urban vendors. These groups had previously viewed one another with some hostility. However, by bringing stakeholders together in a new way, I4ID was able to tap into ideas and capabilities that were possible in the local context, but required some focused attention to emerge. In other words, they were able to tap into a kind of 'latent practice' (Andrews et al., 2013; 2017: 173).

27 Annex, p. 42, para 19.1; p. 29, para 13.1.

28 ibid.

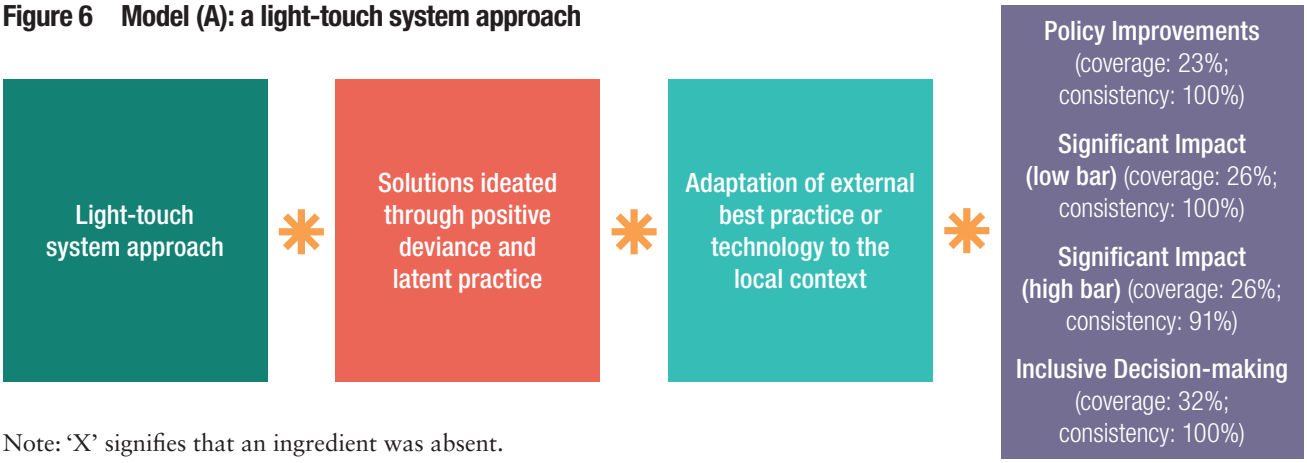
Model (A)
Boolean minimisation was used to dig deeper into the combinations of key causal ingredients. Having tested various models, we found that the only combination with consistently positive results was the presence of a light-touch system approach (2.2), combined with solutions ideated through positive deviance and latent practice (2.5), and a deliberate attempt to adapt external best practice or technology to the local context (2.6).[29] However, while this combination was consistently associated with success across Outcomes 1, 2A and 3, it was somewhat inconsistent with respect to 2B. Nevertheless, the consistency with which this combination contributed to success across most outcomes leads us to recommend it among our core ingredients, outlined at the start of this section.

To illustrate this model, consider again the output in the RIF workstream that focused on facilitating at least one new investor in Iringa.[30] I4ID co-created and prototyped potential solutions with Iringa Regional Government, the Ministry of Regional and Local Government (PO-RALG) and local consultants, through establishing a RIFO. At the outset, the workstream co-ordinator did some mapping to understand the constraints faced by investors in Iringa (2.2). Although the mapping itself was done in a fairly quick and light-touch manner, it nevertheless situated the investment challenge within a more complex network of interrelated issues.

The idea of setting up a RIF team came from the Regional Administrative Secretary in Iringa, who understood the need for investment support (2.5). The intervention also involved external best practice (2.6), insofar as the workstream co-ordinator subsequently brought on board a consultant who was experienced in the United Nations Conference on Trade and Development (UNCTAD) model of investment facilitation.

Models (B) and (C)
When we tested further models (Figure 7), we also found that using human-centred design (2.4) was particularly important for Outcome 3, particularly when combined with rigorous systems analysis (2.1) *or* an attempt to adapt external best practice or technology to the local context (2.6).[31]

Figure 6 Model (A): a light-touch system approach



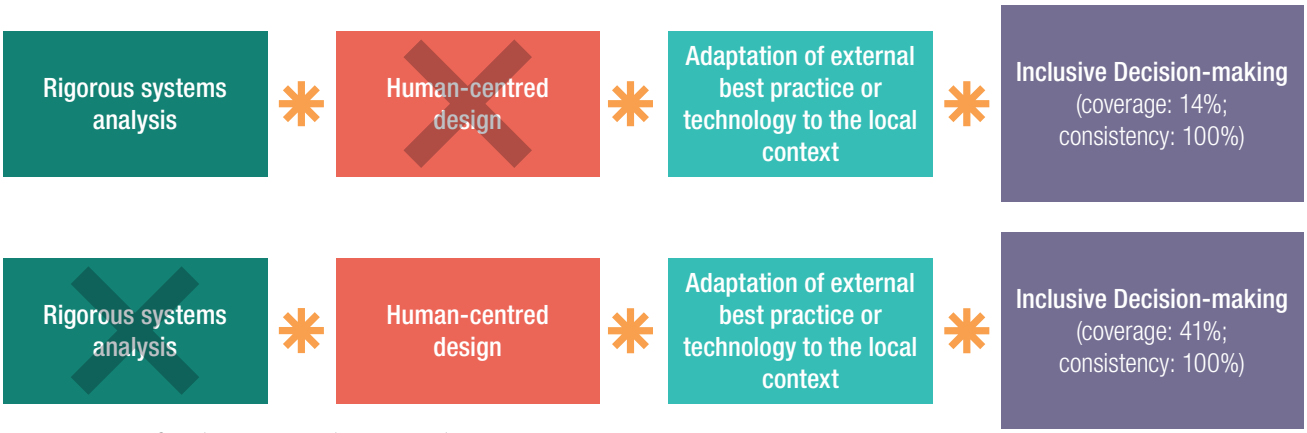
Note: 'X' signifies that an ingredient was absent.

29 Annex, p. 32, para 15.2. When we tested this model, it also included the absence of sub-condition 2.3 in this combination: *problems approached as discrete, one-dimensional issues*. Since 2.3 is absent from all outputs across all the workstreams, we can drop it from our causal model.

30 2.RIF – OP1 (1).

31 Annex, p. 32, para 19.0.

Figure 7 Models (B) and (C): human-centred design for Inclusive Decision-making



Note: 'X' signifies that an ingredient was absent.

To illustrate Model (B), as noted above the USD workstream carried out a fairly intensive system mapping exercise at the outset to understand how land was being used in Kigamboni, and to develop a vision for how land development would be guided in the future (2.1). At the same time, solutions were co-created and prototyped through consultative processes and community feedback, with I4ID engaging with citizen groups and vulnerable communities in Kigamboni, along with specialists from Ardhi University, Kigamboni Municipal Council, the Dar es Salaam Masterplan team and the Ministry of Lands (2.4). I4ID believe they have encouraged a key behavioural change through these efforts, noting that Kigamboni Municipal Council leaders have proactively engaged with and accommodated the interests of a wide, diverse set of stakeholders in a much more transparent and consultative land use planning process, indicating a positive result with respect to Outcome 3.

With respect to Model (C), the IE workstream carried out a study at the outset with its partner CHAVITA to model different parts of the system around disability and education.[32] This was not a particularly rigorous mapping exercise, and it did not involve specific tools such as trend or causal loop analysis, so we felt it did not meet the bar for ingredient 2.1, use of rigorous systems mapping. The workstream placed more emphasis on arriving at possible solutions through community engagement and feedback, and consultations and workshops with political stakeholders, education groups and sign groups (in other words, using human-centred design, 2.4). While the workstream as a whole was driven by local actors, the co-ordinator also drew partly on external best practice, for example taking inspiration from Finnish and Swedish sign practices to inform the TSL harmonisation process (2.6).[33] It seems reasonable to say that this case, and a cluster of other relevant cases in the workstream, [34] have contributed to Outcome 3, inclusive decision-making, given that the Tanzanian government has now formally adopted TSL as the official, harmonised sign language for the country, and has introduced communication regulations to provide sign language translation with the daily news.

These examples illustrate two combinations of ingredients which consistently lead to Outcome 3 success. Given that the common denominator is the use of human-centred design, we have included it among our key ingredients at the start of this section.

32 CHAVITA (Chama cha Viziwi Tanzania) is the Swahili name for the Tanzania Association of the Deaf.

33 38. IE OP2.2 (1).

34 29. IE. OP1 (1); 30. IE OP1 (1); 37. IE OP2.2 IE (2); 38. IE OP2.2 (1).

4.4 Actors

4.4.1 Summary

One way I4ID attempted to go with the grain of local politics, and thus create interventions that had both impact and tractability, was through being highly cognisant of and responsive to the interests and incentives of key actors, and/or enrolling them in its initiatives. We identified four ingredients that a programme like I4ID might use to achieve this, and compared the extent to which these ingredients, and different combinations of them, were present or absent across successful and unsuccessful outcomes. Applying our rules of thumb to these findings, we identified two models with high coverage or consistency rates, which we discuss below. Generalising across these models and drawing on illustrative examples, we would recommend the following key ingredient for future programmers: get buy-in from local non-state actors and implementing agencies, and – especially if organised opposition arises – high-level political leaders.

4.4.2 Models, illustrations and discussion

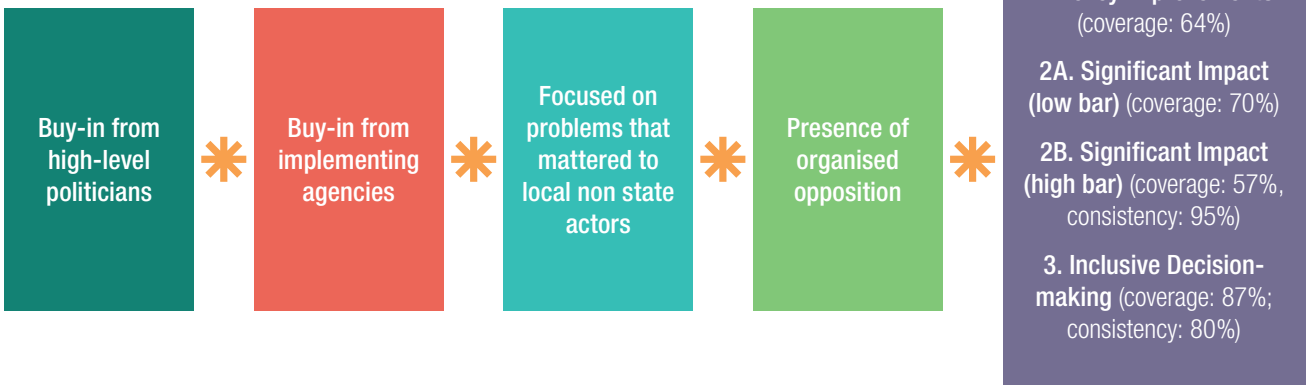
We started by investigating whether any single ingredients were necessary for success. Our results indicated that no single ingredient was necessary, with the exception of buy-in from implementing agencies (3.3), which was necessary for success on Outcome 3, 'Inclusive Decision-making'. Buy-in from local non-state actors was also present in 94% of successful Outcome 2B cases (Significant Impact (high bar)), making it virtually necessary here.[35]

We then carried out Boolean minimisation to generate information on the most important combinations of factors associated with outcome success.

Model (A)

The results were complex (Figure 8), but one model with reasonably high coverage and consistency, especially on Outcome 3 'Inclusive Decision-making', involved a combination of support from high-level political leaders (3.1), implementing agencies (3.3) and non-state actors (3.4), together with the presence of organised opposition (3.5).[36]

Figure 8 Model (A): high coverage and consistency for Inclusive Decision-making



35 Annex, p. 46, para 20.1.

36 Annex, p. 46, para 20.4.

Model (B)

By contrast, where there was no organised opposition (just under half the cases on Outcomes 1 and 2), it was not necessary to have the support of high-level politicians.[37]

To provide an example of this model, in the SWM workstream one case involved a pilot initiative that used geo-tagging to improve fee collection for collecting waste in low-income communities.[38] I4ID deliberately chose to test the initiative in wards where contracts for waste collection had already been granted to its partner, GreenWaste Pro. That made it more difficult for potential blockers to sabotage the issue, and we can reasonably surmise that this was a factor in Outcome 1 and 2 success.

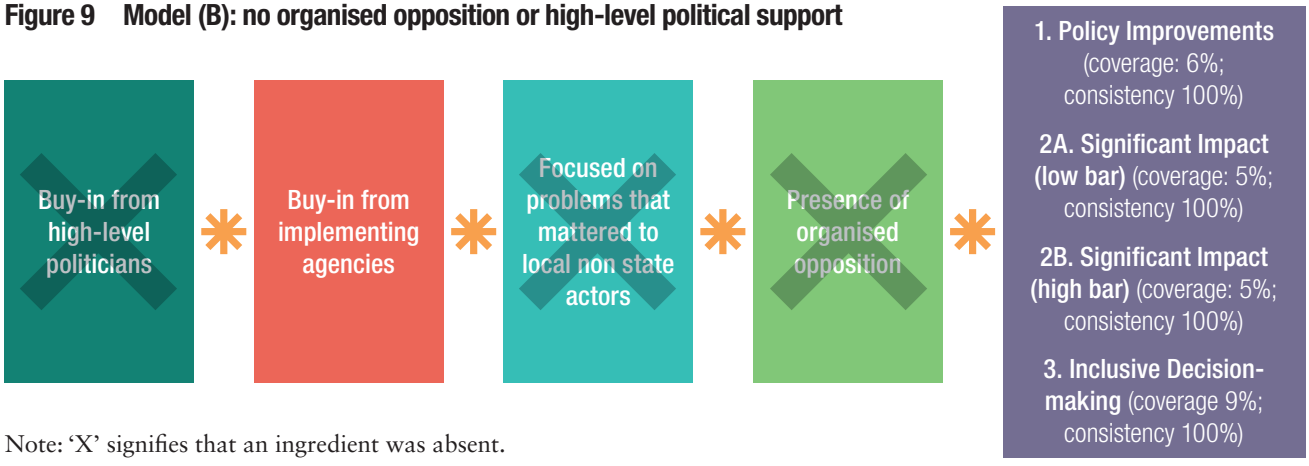
Another illustration comes from the UW workstream. Here, a range of outputs achieved success across all four outcomes, despite only having buy-in from implementing agencies (3.3).[39]

Although this is not reflected formally in our model, the strength of support received from the implementing partner might be relevant here.[40] The UW workstream benefitted from working closely with a particularly committed and effective implementation partner – the Singida Urban Water Supply Authority (SUWASA). SUWASA was well prepared, highly interested in the expected project results and made skilled staff available. It helped drive a network expansion plan based on customer mapping to improve access for under-served households, developed with I4ID support. At a later point, SUWASA convened political leaders and technical stakeholders and secured their agreement for the network expansion plan once implementation was under way.

In the absence of an implementing partner prepared to take the lead in this way, it seems likely that I4ID would have had to generate more explicit buy-in from other actors from the outset.

This interpretation is consistent with some existing arguments in the thinking and working politically literature. For example, it is argued that multi-stakeholder collective action is only likely to succeed if those stakeholders with an interest in the initiative succeeding are able to 'trump' or prevail over those who wish it to fail. Conditions include that 'There are stakeholders with strong incentives to have the collective effort succeed'; 'These stakeholders are well-connected politically'; and 'Leaders of the public organization are skillful

Figure 9 Model (B): no organised opposition or high-level political support



Note: 'X' signifies that an ingredient was absent.

37 Annex, p. 49, para 20.9 and Figures 64–67.

38 1.SWM1.

39 Annex, p. 45–46, paras 20.3–20.7.

40 Our inference is based partly on the findings of a separate study, by one of the authors, looking at I4ID's partnership dynamics: Laws (2020a).

in mobilizing and co-ordinating ... stakeholders in support of the organization's mission' (Levy and Walton, 2013; Levy, 2014).

In the IE workstream, for example, one output was focused on achieving the participation of all major sign language sub-groups in creating a harmonised TSL.[41] I4ID worked in partnership with CHAVITA. This output achieved success across all the outcomes, despite encountering some opposition from different agencies and groups (3.5). Some actors felt CHAVITA did not represent their interests; others had different solutions (for example, focusing on distributing implants rather than harmonising TSL); and some agencies wanted to preserve their unique signing system. Against this, the workstream had high-level political and implementer buy-in, with the Permanent Secretary of the Ministry of Education, Science and Technology (MoEST) expressing particular willingness to engage with the programme, and the Prime Minister launching the TSL Harmonisation Project and calling publicly for relevant authorities to support harmonisation teams (3.1; 3.3). By focusing on meaningful reforms to the education system for deaf people, this intervention can also be said to have focused on problems that mattered to local non-state actors (3.4). These factors were sufficient to overcome whatever opposition the workstream encountered.

To extend our cooking metaphor, we might say that opposition from organised groups is like cooking on a high heat when preparing a pilau. It could be fatal to the success of your dish if your food burns, but having a heavy pan with a thick base (i.e. influential political supporters) can help mitigate this.

4.5 Human resources

There is evidence in the wider literature on adaptive management that finding capable staff with the right kinds of skills and networks is an important success factor when undertaking adaptive programming (see, for example, Booth, 2018). But it is less clear what particular mix of skills, experience and connections are most important. We identified eight conditions that we thought were potential ingredients, and we tested these against I4ID outcomes in various models. Applying our rules of thumb and synthesising across the key findings, we believe the following ingredients for success should be highlighted for future adaptive programme designers:

- ensure that at least one member of the senior leadership team has a track record of successfully delivering an adaptive programme.

And, although the evidence is less conclusive on this point:

- enrol locally based staff, particularly those with strong local networks or membership (past or present) in relevant national organisations.

4.5.1 Models, illustrations and discussion

We began by testing whether there were any necessary and/or sufficient conditions for success by looking at whether any were present or absent across all successful or unsuccessful outcomes. We found that all cases involved team members who were resident in Tanzania (4.2), at least one of whom (the team leader) had a strong track record in delivering successful adaptive programmes (4.8).[42] We can infer from this that these may be important factors in success, and programme designers looking to replicate a programme like I4ID would be well-advised to include them in their own design. However, because they are found in all cases, both successful and unsuccessful, we must also concede that a programme could fail despite including them, depending on the other ingredients it featured.

41 38. IE OP2.2 (1).

42 Annex, page 61, para 24.1.

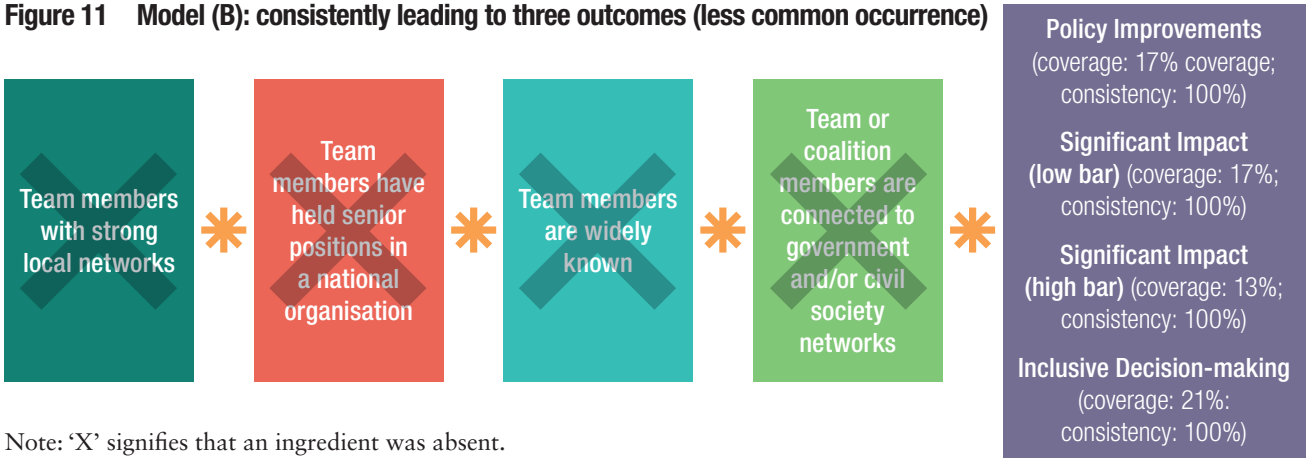
Model (A)
We then tested various combinations of those four conditions that remained once we removed those which were always either present or absent (Figure 10). It is important to note here that there were no combinations that satisfied our rules of thumb with respect to coverage and consistency. With that caveat in mind, the most consistently successful combination of ingredients, at least when it came to the first three outcomes, was to have team members or coalition members with strong local networks, who had previously or concurrently held senior positions in a national organisation, and who were widely known.[43]

It is important to note that, for every Outcome, the successful cases included under this combination are almost all clustered in one workstream: IE. I4ID's key partner in this workstream, CHAVITA, has strong local networks in the disability community (4.1); is a national organisation (4.2); and the original workstream co-ordinator was widely known among disability organisations (4.4).

Model (B)
A less common but still successful combination across all outcomes was for all these ingredients, as well as strong connections to government and civil society networks, to be absent (Figure 11).[44]

Figure 10 Model (A): consistently leading to three outcomes

Figure 11 Model (B): consistently leading to three outcomes (less common occurrence)



Note: 'X' signifies that an ingredient was absent.

43 Annex, p. 60, para 24.3.

44 Annex, p. 60, para 24.4.

However, as with the above combination, we find the successful cases under this combination uniquely associated with one workstream: this time UW. It may seem counter-intuitive to identify successful outcomes with the *absence* of a string of ingredients. However, this finding directs our attention to factors captured under other groups of conditions. In this case, it seems to lend further weight to our contention, explored in our discussion of actors, that the key to the success of the UW workstream was the strength of the partnership with SUWASA. This relationship was not the product of any of the pre-existing networks, relationships, or national membership connections that I4ID staff brought with them into the project, and so the success of the workstream was not undermined by the absence of these conditions.

As we explored different combinations of the presence and absence of these four conditions across the different outcomes, we found the amount of inconsistency increasing.[45] In other words, it became harder to discern any meaningful patterns. Moreover, with respect to Outcome 3, 'Inclusive decision-making', the majority of cases cluster in combinations that do not consistently represent either success or failure. This provides a clue that the ingredients under this heading may not be critical to whether many interventions did or did not contribute to this outcome, just as spices are not critical to how filling a pilau is. Nevertheless, being mindful of our rules of thumb, the models we have discussed indicate it would be wise for future programmers to include the two core ingredients highlighted at the start of the section.

4.6 Learning

4.6.1 Summary

Adaptive programmes start from a position of epistemic humility. They all acknowledge uncertainty about how to achieve their outcomes and engage in a deliberate process of experimentation and learning to build evidence and approach solutions. However, there are many ways of going about these learning processes, and there is little clear sense in the wider literature of what kind of tools and approaches are more or less suitable in different contexts and intervention areas (Laws and Marquette, 2018). Applying our rules of thumb to the findings from our models and illustrations, we recommend that future programmers should:

- take a sequential approach to testing and learning
- have a credible plan in advance for taking interventions to scale.

Although the evidence is a little less strong on this point, it is probably also wise to start interventions with at least a loose theory of change, and preferably a set of clear hypotheses.

4.6.2 Models, illustrations and discussion

When we scanned our results, we noted that I4ID rarely tested multiple hypotheses simultaneously (5.2), preferring to place one bet on a potential solution to a problem, and adjust course if necessary once implementation was under way.[46] In the SWM workstream, for example, one output was geared towards expanding SWM services in Ukonga Ward.[47] Rather than experimenting with a range of models simultaneously, I4ID tested one approach, based on a mobile transfer station using a tractor and trailer. The initial experiment was designed to provide proof of concept before being scaled up to other wards. After the successful pilot, I4ID prompted replication by identifying other suitable companies.

Even where, as with other interventions in this workstream, multiple initiatives were trialled simultaneously, these addressed different *bits* of the problem of SWM: they were not alternative solutions to the same problem. Given that this sequential approach was present in the vast

45 Annex, p. 60, para 24.6.

46 Annex, p. 65, para 25.0–25.1.

47 8. SWM6.

majority of our cases, we are unfortunately unable to test, in any rigorous way, for the relative strengths of this compared to a parallel approach. However, in keeping with our rules of thumb, we can still recommend a sequential method as a sensible place to start for programmers looking to replicate I4ID's achievements.

We also noted that few cases involved the systematic analysis of results against initial hypotheses (5.5), with workstreams more commonly engaging in what we would describe as 'purposive muddling'.[48] In other words, a process of testing and learning that is still deliberate, somewhat intentional with respect to testing initial hypotheses, and supported by flexible programming architecture, but with a mostly ad hoc and informal approach to the timing, regularity and format for reflection and analysis. Across several cases in the RIF workstream, for example, the trajectory of interventions was largely shaped by the political intuitions of the workstream co-ordinators, the I4ID team leader and delivery partners in the RIFO, rather than by reference to progress against indicators in the workplan, or through structured processes of reflection and evaluation. Efforts were made, at times, to create more formalised methods for tracking progress against initial hypotheses. For example, a stakeholder engagement document was created and shared across the intervention team to keep track of the various relationships being cultivated. But the workstream co-ordinator found it hard to document and explain all the small pivots and course corrections that happened from week to week in the course of trying to build and maintain these relationships, because so much of it was based on instinct.

To revert back to our central metaphor, some chefs might argue that the key to a consistently successful pilau is tasting the food at set intervals in the cooking process and marking it against a clear set of predetermined criteria. Another chef might prefer to take a more relaxed approach, tasting the food now and then as they see fit, and measuring it against a looser and more intuitive sense of the qualities they are looking for. This kind of chef might argue that, by insisting on documenting everything in granular detail against a clear list of criteria, you stifle the creativity required to produce excellent results.

Returning to our dataset, to test combinations of ingredients we removed from our analysis conditions that are rarely present (5.2 and 5.5) and tested various combinations of the remaining five. Although the results for these combinations were extremely complex, two models consistently delivered positive results across all our outcomes, and a third was associated with relatively high success rates in Outcomes 2B and 3.

Model (A)

The first was the absence of rigorous hypothesis testing (5.4), with the presence of loose hypothesis formation (5.1), conditions for testing established (5.3), experimental, iterative steps progressively employed (5.6), and a clear plan for scale-up (5.7) (see Figure 12).[49]

These cases are all clustered in the RIF workstream. As noted, as a general rule this workstream followed a process of purposive muddling rather than structured experimentation. There were no fixed points or processes for testing and pivoting, and it is not clear whether the outputs involved much deliberate testing of hypotheses. What these consistently positive cases suggest is that it is not necessary for adaptive programmes to follow a very formalised testing process if other aspects of the learning process are established clearly. This is what distinguishes *purposive* muddling from muddling, and these cases provide an important counterweight to concerns that 'being adaptive'

48 Annex, p. 65, para 25.1. This term was popularised by Matt Andrews in the context of developing the PDIA methodology at Harvard. See Andrews et al. (2013). However, the idea of 'muddling' has a longer pedigree in political science, with the earliest reference of which we are aware appearing in Lindblom (1959). Our definition departs from Andrews' and Lindblom's by emphasising the lack of formal structure around the timing, regularity and format for reflection and analysis.

49 Annex. p. 74, para. 29.4 and Figures 96–99, Box 11011.

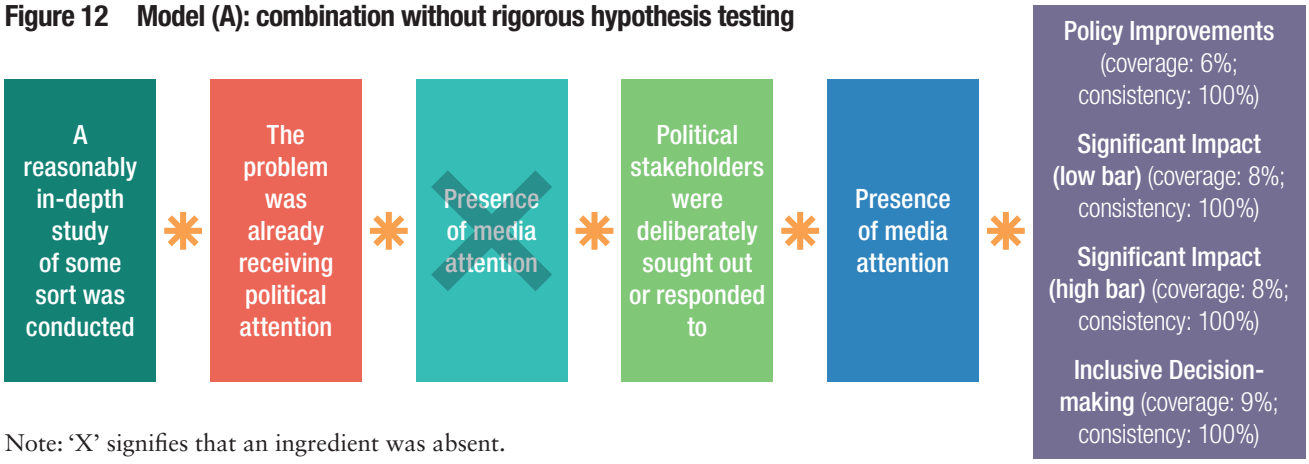
is often a fancy shorthand for 'making it up as you go along' (Ramalingam et al., 2019: 3).[50]

Note that, although this combination consistently led to success, it should be stressed that it covered only three cases. Loose hypothesis formation (5.1) by itself was not consistently associated with success, and there were a few successful cases where it was coded absent. That said, those cases were found uniquely in the UW workstream.[51] And while it may be true that some of that workstream's specific outputs did not involve this condition, it would be inaccurate to say that the workstream as a whole did not have a loose theory of change. Consequently, it may be advisable to treat the loose hypothesis ingredient as more or less necessary. As per our rules of thumb, we have included it among our core ingredients, summarised above.

Model (B)

The second model that consistently contributed to success across all four outcomes was a plan for scale-up (5.7), combined with the absence of all the other conditions (Figure 13). These positive cases were all clustered in the UW workstream, where there were plans from the outset to scale

Figure 12 Model (A): combination without rigorous hypothesis testing



Note: 'X' signifies that an ingredient was absent.

Figure 13 Model (B): combination including a clear plan for scale-up

Note: 'X' signifies that an ingredient was absent.

50 Another combination, which is worth highlighting but which did not fall under our rules of thumb for identifying key findings, was the presence of hypothesis formation (5.1), with the absence of systematic analysis against success and failures (5.5), and the absence of a clear plan for scale-up (5.7). This combination covered 16% of cases with a positive Outcome 1 and 10% of cases with a positive Outcome 2A. Annex, p. 68, para 27.1.

51 See Annex, pp. 68–71, Figures 88–91, Box. 0001.

the approach to network expansion being taken in Singida to at least one other municipality.[52]

Model (C)
A final model, which tested a slightly different combination of ingredients, delivering relatively high coverage on Outcomes 2B and 3, combined a plan for scale-up (5.7) with the absence of multiple hypothesis testing (5.2) and the absence of systematic analysis of successes, setbacks and failures (5.5) (Figure 14). In addition to contributing to success across I4ID's more challenging outcomes, it is also noteworthy that the successful cases with this combination of ingredients are spread quite widely across I4ID's workstreams.[53]

Yet note that behind this presence of a plan for scale-up there arguably lay another condition, namely the reluctance of the team leader to fund any initiative that did not have a plan he found credible. Not without controversy, there were various I4ID initiatives – most notably, perhaps, one to promote a home-grown version of an improved sunflower seed – that fell by the wayside for this reason. Consequently, it is difficult to say with certainty that initiatives would not have worked without a credible route to scale articulated in advance. Despite this uncertainty, and as with other necessary but perhaps incidental ingredients, we include it in our advice to future programmers.

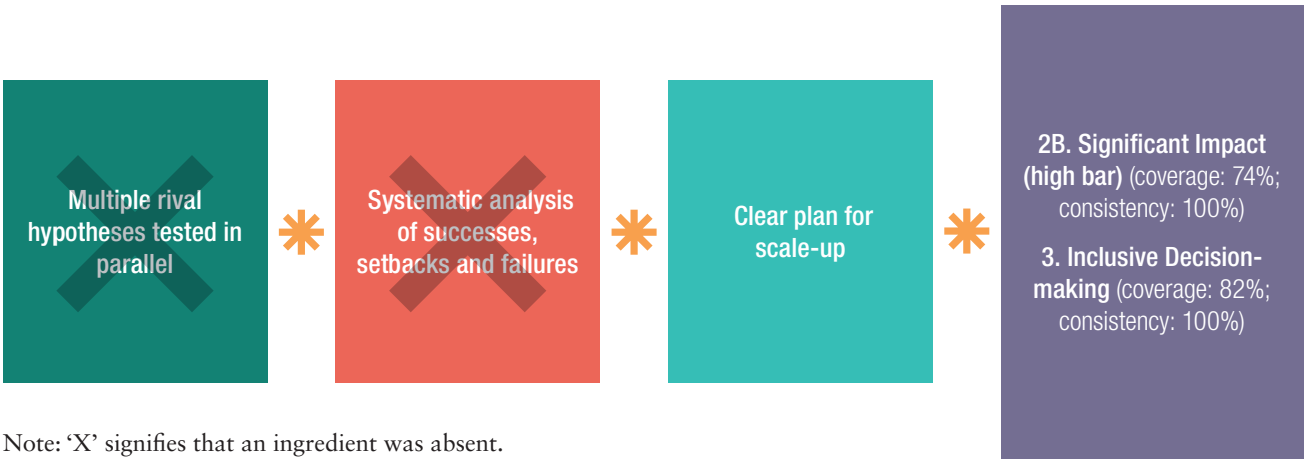
4.7 Funding

4.7.1 Summary

All aid programmes, adaptive or otherwise, involve the management and disbursement of funds, yet it is often stated that the common imperative among donor agencies to spend in a predictable manner is problematic for adaptive and politically smart programmes, which need to be more flexible and strategic in their spending (Wild and Foresti, 2011; Rocha Menocal, 2014; Wild et al., 2015; 2017; Teskey and Tyrrel, 2017). It has also been noted that politically informed work has seldom been associated with large financial investments, but it does tend to involve higher managerial and administrative overheads (Piron et al., 2016: 35). This can present challenges for donors, where efficiency is often equated with programmes that can disburse large amounts of money with relatively small administrative costs (Laws and Valters, 2021).

How, then, can funding be managed in ways that better support politically smart and adaptive programming?

Figure 14 Model (C): a combination for the more challenging outcomes



Note: 'X' signifies that an ingredient was absent.

52 Annex, p. 74, para 29.4 and Box 00001 in Figures 96–99.

53 Successful cases for Outcome 2B with this combination were found in SWM, UW, MHM and RIF. Successful cases for Outcome 3 with this combination were found in UW, RIF and USD. See Annex, para 28.2 on p. 71, Box 001 on Figures 94 and 95.

Following our rules of thumb, we wish to highlight the following ingredients as advice to future programmers:

- Use funds in a strategic way.
- Be prepared to provide funds for pilots, small-scale infrastructure, technical assistance or core funding.
- Use funds for brokering and convening, though not necessarily as the principal approach.

4.7.2 Models, illustrations and discussion

Our models showed that using funds in a strategic way (6.2) and having sufficient budget (6.5) were (almost) necessary for success. However, we do not read too much into the budget ingredient, since having sufficient funding is arguably necessary by definition for the success of any development intervention.[54]

Our findings for ingredients are somewhat more interesting. A core part of I4ID's self-understanding was its commitment to keeping 'money off the table' when working with partners (6.3): moving away from the common approach of providing grants to civil society organisations (CSOs) or making large capital investments, and instead focusing on facilitation, brokering and convening. This was partly an effort to avoid contributing to a perceived culture of aid dependency in Tanzania, and partly a wager on the greater effectiveness and sustainability of self-motivated and self-reliant reform efforts.

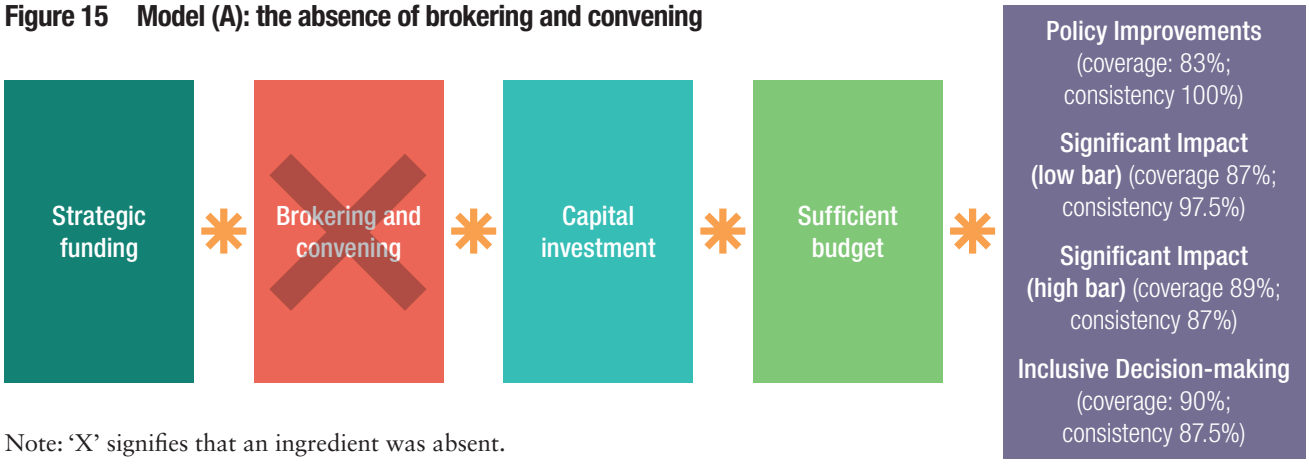
However, our results reveal that, despite the aspiration to keep money off the table, most of I4ID's successes also involved some form of capital investment, that is to say funds for pilots, small-scale infrastructure, technical assistance or core funding, in response to a request from a partner (6.4).[55]

In fact, when we tested various combinations for the presence and absence of ingredients, we found that, in most successful cases, convening and brokering was not the principal approach (6.3). And even in the small number of cases – limited to the SF workstream – where it was, success was still accompanied by some form of capital investment (Figure 15).[56]

Model (A)

It is important not to misinterpret this. Brokering and convening were defined quite narrowly, so they are only coded as 'present' when they were the *principal* approach in a particular intervention. However, capital investment was defined expansively, so that it is coded as present whenever it occurred. The truth is that almost all I4ID workstreams involved some funding for convening and brokering, so although it is clearly not necessary for it to be the principal approach, that does not mean it is not necessary.

Figure 15 Model (A): the absence of brokering and convening



Note: 'X' signifies that an ingredient was absent.

54 Annex, p. 85, para 33.0.

55 Annex, p. 91, para 37.1.

56 Annex, pp. 85–87, paras 34–35; Figures 112–115, bottom half of the charts, outside the central rectangle.

To provide a few illustrations. In the SWM workstream, to support changes in the procedures for contracting and regulating SWM in Ilala Municipality, I4ID convened several technical roundtable meetings between Municipal health officers, Ward Executive Officers, service providers or small and medium enterprises (SMEs) and local residents, with a view to exploring the drivers of and potential solutions to the problem of dumping and poor waste collection.

However, these initiatives also involved providing funding for pilots and technical mapping exercises, among other things.

In the IE workstream, meanwhile, I4ID's main partner CHAVITA struggled with cash flow, partly as a result of I4ID's insistence on only reimbursing expenses rather than providing direct funding. I4ID eventually agreed to CHAVITA's request for a core grant, reasoning that having them on board was important enough to justify departing from the money off the table philosophy.

Unfortunately, the QCA does not provide much insight into the best sequencing of these different ingredients. However, a deeper examination of individual cases suggests that it may be advisable for an adaptive governance programme like I4ID to start with brokering and convening and maintain a presumption in favour of keeping money off the table, particularly if they understand the principal blockage they are working on to involve a coordination challenge. But once implementation gets under way, I4ID's experience suggests it would be unhelpful to adhere rigidly to this modality. Indeed, it would be hard to classify as flexible and adaptive any programme, successful or otherwise, that treated something like money off the table as a red line rather than a rule of thumb.

It is also instructive to look at the UWV workstream in this context, which collaborated with the Mwanza City Council, municipal councils and urban vendors' groups to co-create solutions for safer vending spaces. I4ID was initially reluctant to agree to requests from local stakeholders for investment in market infrastructure, on the grounds that it would divert too many resources towards fixed infrastructure spending and make it difficult to maintain a focus on gender outcomes. However, for as long as it stuck to this strategy, the workstream found it challenging to make any headway. Once I4ID agreed to contribute some funds to market infrastructure, it was able to use these investments as levers to promote gender-friendly institutional reform, as well as integrating specific targeted benefits for women traders in the design of the new physical structures.

For these reasons, and as stated in the summary at the start of this section, in contexts like Tanzania we advise future programmers to use funds strategically, for both capital investments and brokering and convening.

4.8 Capacity and innovation

4.8.1 Summary

To address the often complex and political challenges that its workstreams faced, I4ID typically aimed to introduce or build different kinds of capacity in partner organisations and other stakeholders. Depending on the type of problem, the programme sometimes tried to do this by searching for, or trying to bring about, innovative behaviour or technologies. In other cases, workstreams focused on existing solutions and tried to improve upon them.

Following our rules of thumb, we find that future programmers should:

- introduce and/or build capacity for an innovative technology; *and/or*
- introduce and/or build capacity for innovative behaviour; *and/or*
- 'fine-tune' an existing system, model or technology, as needs arise.

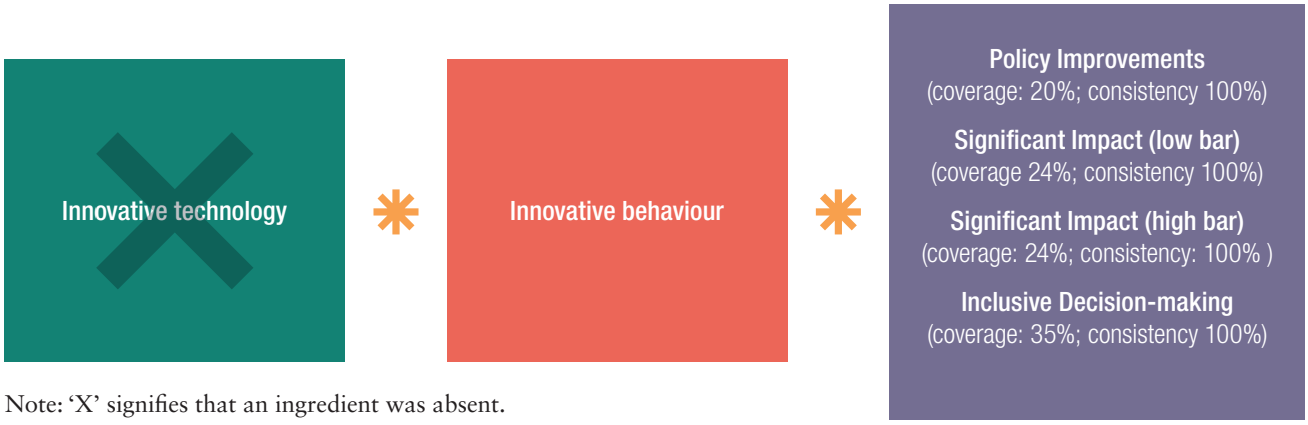
4.8.2 Models, illustrations and discussion

Superset analysis for these ingredients showed that none of them, by itself, was necessary for success.

Model (A)

When we ran Boolean tests on the three-condition model, we found none of the cases that was successful on either of the first two outcomes were associated uniquely with any particular combination (Figure 16). The only combination that stayed consistently positive across the four outcomes, meanwhile, is the absence of

Figure 16 Model (A): omitting innovative technology; including innovative behaviour



Note: 'X' signifies that an ingredient was absent.

innovative technology (7.1) with the presence of innovative behaviour (7.2).[57]

The MHM workstream, for example, tried to catalyse changes in the market system for menstrual products by encouraging new, more collaborative behaviour among competitor companies. In one output case, this culminated in seven companies overcoming competitive tensions and committing to joint marketing strategies and activities for market growth for the first time.[58] The key to success in this case was innovative behaviour, rather than new technology.

However, this model had quite low coverage, and other models that involved the *presence* of innovative technology were significant. In the SWM workstream, for example, I4ID provided funding for innovative approaches to recycling, including the use of black soldier flies to turn organic waste and manure into high-protein fish feed.[59] In the UW workstream, the programme helped its partner SUWASA build an operational GIS platform and customer management software.[60]

It is also worth noting that almost half of the successful cases in 'Policy Improvement' and 'Significant Impact (low bar)' *did not* involve innovative behaviour *or* technology, with I4ID working to improve an existing solution, model or approach.[61] One case in the IE workstream, for example, focused on introducing relatively small but important improvements to the existing quality assurance framework for IE.[62]

In general, though, there seems to be quite a high concentration of contradictory combinations, especially for 2B 'Significant Impact (high bar)' and 3 'Inclusive Decision-making'.[63] One might infer from this that the ingredients under this heading were not critical to I4ID's success. However, we have a more cautious interpretation. Returning to our pilau analogy, one can imagine a situation in which one has employed a chef who can cook a biryiani, but not a pilau. You have a choice of helping them to improvise on this existing knowledge, apprenticing them to an expert pilau chef, or teaching them how to microwave pilau ready meals. Choosing one of these strategies is preferable to doing nothing, and the choice should be determined by the specific needs of the case.

This reasoning explains the 'either-or' ingredients summary presented at the head of this section.

57 Annex, p. 94, para 39.0.

58 22. MHM OP2.1 (2)

59 4. SWM2

60 12. UW1

61 Annex, p. 95, para 38.1.

62 34. IE OP1.IE(5)

63 Annex, p. 97, Figures 130–131.

5 Conclusions and recommendations

In this report, we hope to have shown policy-makers, programme designers and programme commissioners in donor agencies interested in starting adaptive, politically smart governance programmes in Tanzania or countries like it the ingredients most meaningfully associated with a track record of success in I4ID: the core ingredients. We hope to have demonstrated their importance in a way that has greater rigour and transparency than previous analyses. As far as we are aware, this is the first time QCA has been used to investigate the success factors associated with these approaches.

5.1 The core ingredients

What, then, *are* the core ingredients for adaptive governance programmes in Tanzania, and in countries like Tanzania? And what plausible additional inferences can we draw, from our analysis, to help steer the designers, commissioners and leaders of the next generation of these programmes?

5.1.1 Politics

Programmes should focus on addressing problems that are already receiving a high level of political attention, and around which supportive political stakeholders or champions can be assembled. It is also important that the political support of the funding organisation is maintained.

Our results indicate that a light-touch approach to PEA may be sufficient. However, on the basis of our broader understanding of how PEA was used in the programme, we would also advise that in-depth analysis be applied more rigorously at a programme-wide level at relevant junctures, such as major changes in the political context, and when working on particularly entrenched or complex political problems.

5.1.2 Design

Rigorous systems research should be regarded as a potentially useful optional extra in a core ingredient combination, but not central to success. A light-touch approach to system analysis and mapping is usually enough to deliver good results, particularly when combined with positive deviance, co-creation or leveraging external best practice.

Programmes that are focused, at an outcome level, on bringing about more inclusive decision-making, planning or policy processes should incorporate human-centred design principles. It is also wise to start interventions with a credible plan for scale-up.

5.1.3 Actors

While it may be advisable to look for pathways to change that are unlikely to meet resistance from organised interests, this kind of opposition can be overcome provided the initiative has the support of high-level and influential political leaders, implementing agencies and non-state actors.

In the absence of organised opposition, success can be achieved with a smaller range of supporting stakeholders.

5.1.4 Human resources

It helps to hire locally based staff, particularly those with strong local networks or membership in relevant national organisations. This appears to be more critical than their experience in technical analytical methods.

It is not essential that staff have a background in adaptive management or applied PEA, provided at least one member of the programme leadership has a track record of successfully delivering these approaches and methods.

5.1.5 Learning

The right choice of learning strategy depends on the nature of the problem at hand, but taking a sequential approach is a reliable route to success.[64]

Success does not depend on having a very formal process in place for testing, provided in most cases that there is at least a loose theory of change and set of hypotheses that are frequently referred to and re-evaluated.

5.1.6 Funding

Our data does not demonstrate clearly that brokering and convening (i.e. keeping money off the table) is more closely associated with success than the provision of capital investments or core grants, or vice-versa. As such, we would advise that, while future programmers may choose to begin their engagement with brokering and convening, they should be prepared to provide funds for pilots, small-scale infrastructure, technical assistance or core funding.

5.1.7 Capacity and innovation

Issue-based adaptive governance programmes should not set themselves up as either innovation or 'fine-tuner' programmes from the outset, but should be flexible, and respond with appropriate methods and capacity-building strategies as needs arise.

5.2 A transferrable recipe?

Despite using QCA, we should be a little cautious about the degree of wider relevance that we claim for our findings. We can say with real confidence that these combinations of ingredients worked well in this place at this time, but as we expand beyond that context, we expect diminishing circles of confidence in the transferability of this recipe.

That said, we think it is reasonable to predict that a programme with these core ingredients would stand a *good chance* of succeeding not only in Tanzania, recent post-election changes notwithstanding, but also in other relatively stable lower-middle-income countries, with constrained political space and low to moderate levels of state, private sector and civil society capacity.[65]

There is also little reason to think that a programme with these core ingredients would not work in contexts with wider political space and greater capacity, although there might arguably be less need for such a programme, and/or it might want to be more adventurous, especially in its political approach.

A programme with these core ingredients *might* also work in contexts with significantly less capacity, though a more arduous search for credible partners, with a higher risk of failure, would, perhaps, be involved. The same observation applies to less stable political contexts, where risks are generally higher across the board.

By contrast, it is *hard to imagine* a programme with I4ID's DNA working in contexts with *extremely narrow* political space, as civil society activity is likely to be difficult to find, and mistrust of quasi-independent development initiatives entrenched.

5.3 What's new about our findings?

None of these ingredients themselves, or their successful combinations, are likely to come as a complete surprise to anyone familiar with the literature on adaptive programming and on thinking and working politically. However, we think our study also has some less obvious lessons to add to the conversation. With respect to adaptive rigour and adapting spending, there are two findings from our study which, while they may not fully overturn received wisdom, do provide some additional nuance.

5.3.1 Adaptive rigour

Large bureaucracies and development organisations can have low tolerance for experimentation and learning, and adaptive management is sometimes viewed as an excuse for 'making things up as you go along' (Ramalingam

64 This means trailing one approach or potential solution to a problem and adapting it over time.

65 An important caveat is that the success of programmes like this is heavily dependent on the quality of their staff. Although we have tested some generic categories for human resources, we have not been able to code individuals' unique personalities, passions, talents and foibles.

et al., 2019: 3). This has led, in recent years, to investment in building greater rigour in the systems and processes that support these approaches, including formal tools and guidance for PEA and systems analysis, and in strengthening the quality of monitoring, evaluation and learning data and systems.[66] But I4ID has demonstrated that it can achieve success with a somewhat looser and more informal approach in each of these areas – largely captured by what we have referred to as 'purposive muddling'. In-depth PEA, rigorous systems research and analysis, and systematic analysis of results against initial hypotheses figured as potentially useful optional extras, rather than central ingredients to success. As one of our reviewers underlined, a plausible inference from these findings is that it may be more valuable to do frequent, 'everyday'-type PEA and systems thinking as part of an engagement, rather than spend a lot of time and money upfront on formal studies and mapping (see, for example, Marquette et al., 2016).

However, we would reiterate the need to view these results with some caution, noting in particular that our findings are still consistent with the idea that achieving *sustained* progress on more complex and entrenched governance problems is likely to require more in-depth analysis and greater upfront and ongoing investment in these areas than was the norm in I4ID. Moreover, in other programmes, systematic forms of analysis and evaluation have played an important accountability function and helped to ensure a supportive donor authorising environment, something we flagged as a key ingredient in our discussion of the 'actors' important to I4ID's success (see, for example, Laws, 2018).

5.3.2 Adaptive spending

A core part of I4ID's self-understanding was its commitment to keeping 'money off the table' when working with partners, partly to try to demonstrate that complex development problems can be resolved by bringing diverse stakeholders together to address co-ordination and collective action problems. Most of the workstream co-ordinators and senior team leaders referred regularly to this as a guiding principle during our conversations and interviews, although some regarded it as an impediment, rather than an ingredient of success.

This philosophy fits with a broader narrative in the adaptive TWP literature, which stresses the importance of brokering, convening and building relationships with local partners in which aid money is not the key motivating factor. For example, fostering relationships and alliances around common interests were critical activities in all of the case studies documented in a widely cited study by Booth and Unsworth (2014). Laric and Waddell (2016) discuss how a DFID-funded[67] facility in Nepal, the Centre for Inclusive Growth (CIG), helped overcome constraints to economic growth and private sector investment through politically smart deal-brokering between multinational corporations, private investors and the government.

However, our results reveal that, despite the aspiration and guiding philosophy about keeping money off the table, most of I4ID's successes in practice also involved the strategic investment of programme funds. Convening and brokering was rarely the principal approach in most successful cases, and even in the small number of cases where it was, success was still accompanied by some form of capital investment, and even core funding to a partner organisation in one workstream.

Again, however, we encourage the reader to interpret the wider relevance of these findings with subtlety, and with due consideration to the broader context. Booth and Unsworth (2014: 19) point out that none of the programmes they studied was under pressure to meet particular spending or results targets, and had the luxury of relatively long-term funding commitments from their donors. CIG was operating in a context where the investment climate had stagnated for decades, so we might reasonably infer that the expectations on the programme were

66 See, for example, https://usaidlearninglab.org/cla-toolkit; www.odi.org/projects/2918-global-learning-adaptive-management-initiative-glam.

67 The UK's Department for International Development, which merged with the Foreign and Commonwealth Office in 2020 to become FCDO.

manageable. By contrast, having had a rather turbulent start I4ID was under considerable pressure to generate results for DFID quickly, from its second year onwards. That being the case, it appears that the senior programme leadership, not unreasonably, were at times pragmatic about using funds to speed up reform efforts or unlock progress, in situations where a more patient convening process may well have delivered the same or better results, but perhaps with a longer timeframe and fewer guarantees.

5.3.3 How useful was the QCA approach?

As mentioned in the introduction, the evidence base on adaptive programmes largely comprises anecdotal success stories. Case studies often provide a lot of detail on individual reform episodes and programming processes, but tend not to use rigorous methods to explore how and why combinations of causal factors did, or did not, contribute to programme outcomes. We specifically chose QCA to try to plug these methodological and evidence gaps.

How well, then, did the method hold up, and can we be sure that we have achieved a higher level of rigour than is typical?

Our first observation is that we have been able to provide precise data about the level of success of specific ingredients. Referring to the example above, some I4ID staff took issue with the 'money off the table' principle and thought providing capital investment was important to success in the Tanzanian context. In a conventional study, this may have remained at the level of an assertion, opinion or hunch. But through QCA's rigour, we have been able to demonstrate precisely how commonly this ingredient was associated with success, and draw conclusions therefrom.

Indeed, proceeding in this very systematic way we have been able to highlight ingredients that were present in all or a very high percentage of successful outcomes, and/or that were part of combinations that were consistently associated with success. As such, we have arrived at a 'recipe for success', outlined above, that, with a couple of caveats, we can be confident will deliver good results in Tanzania and countries like it.

However, we should also be frank about some of the shortcomings of the method and our application of it. Two issues seem particularly important.

First, for the purpose of reaching strong conclusions about the necessity and/or sufficiency of the ingredients for successful outcomes, there was less variation in our data than we would have liked. Although we tested an unusually large and complex dataset, some ingredients were present across all or almost all cases. Consequently, some of our claims about necessity have to be taken with a pinch of salt. From the point of view of building an evidence base – if not of avoiding risk – it would be helpful if future programmers could experiment with leaving some of these ingredients out. Moreover, as an adaptive programme I4ID tended at an early stage to drop or rework those outputs and workstreams that were not delivering results quickly enough, meaning that the number of failed cases was low, especially on Outcomes 1 and 2A. Partly for this reason, we do not make strong claims about the sufficiency of various ingredients.

Second, we decided not to measure the success of cases in contributing to the fourth outcome in I4ID's Results Framework:

> Significant instances where democratic institutions involved in the programme demonstrate sustained or repeated behaviour that is consistent with more inclusive decision-making, planning or policy process.

To reiterate, we opted not to include this outcome because, for the most part, it is too early to assess fairly whether outputs have contributed to it. While this decision is methodologically defensible, it does somewhat limit the strength of our conclusions. Put simply, we cannot say with confidence that our recipe for success will contribute to the sustained adoption and scale-up of inclusive institutional change.

Both these shortcomings present challenges that future researchers might choose to take on. It is therefore fortunate that one of the principal strengths of QCA is its transparency and replicability. Readers will note that we have made available all our findings online, should they wish to test our recipe against other programming examples in other countries, or rerun our tests in Tanzania in later years, once the important policy processes that I4ID has contributed to have had more time to evolve.

6 References

Andreoni, A. (2018) *Anti-corruption in Tanzania: a political settlements analysis*. Anti-Corruption Evidence Working Paper 1. London: School of Oriental and African Studies (https://ace.soas.ac.uk/working-paper-1).

Andrews, M., Pritchett, L. and Woolcock, M. (2013) 'Escaping capability traps through problem-driven iterative adaptation (PDIA)' *World Development* 51: 234–244(www.cgdev.org/publication/escaping-capability-traps-through-problem-driven-iterative-adaptation-pdia-working-paper).

Andrews, M., Pritchett, L. and Woolcock, M. (2017) *Building state capability: evidence, analysis, action*. Oxford: Oxford University Press.

Befani, B. (2016) *Pathways to change: evaluating development interventions with Qualitative Comparative Analysis (QCA)*. Expertgruppen för biståndsanalys (EBA, Expert Group for Development Analysis) (https://eba.se/wp-content/uploads/2016/07/QCA_BarbaraBefani-201605.pdf).

Booth, D. (2018) *Incubating policy for economic transformation: lessons from Nepal*. ODI Report. London: ODI (www.odi.org/publications/11090-incubating-policy-economic-transformation-lessons-nepal).

Booth, D. and Chambers, V. (2014) *The SAVI programme in Nigeria: towards politically smart, locally led development*. London: ODI (www.odi.org/publications/8876-politically-smart-nigeria).

Booth, D. and Unsworth, S. (2014) *Politically smart, locally led development*. London: ODI (www.odi.org/publications/8800-politically-smart-locally-led-development).

Burns, D. and Worsley, S. (2015) *Navigating complexity in international development: facilitating sustainable change at scale*. Rugby: Practical Action Publishing.

Carothers, T. and De Gramont, D. (2013) *Development aid confronts politics: the almost revolution*. Washington DC: Brookings Institution Press.

Dasandi, N., Laws, E., Marquette, H. and Robinson, M. (2019) 'What does the evidence tell us about "thinking and working politically" in development assistance?' *Politics and Governance* 7(2): 155–168 (www.odi.org/publications/11374-what-does-evidence-tell-us-about-thinking-and-working-politically-development-assistance).

Dasandi, N., Marquette, H. and Robinson, M. (2016) *Thinking and working politically: from theory building to building an evidence base*. Birmingham: University of Birmingham.

Eriksen, S.S. (2018) *Tanzania: a political economy analysis. Report commissioned by the Norwegian Ministry of Foreign Affairs*. Oslo: Norwegian Institute of International Affairs.

Faustino, J. and Booth, D. (2014) *Development entrepreneurship: how donors and leaders can foster institutional change*. Working Politically in Practice Series. London: ODI (www.odi.org/publications/9118-development-entrepreneurship-how-donors-and-leaders-can-foster-institutional-change).

Green, D. (2016) 'Where have we got to on adaptive learning, thinking and working politically, doing development differently etc? Getting beyond the people's front of Judea'. Oxfam from Poverty to Power Blog (https://oxfamblogs.org/fp2p/where-have-we-got-to-on-adaptive-learning-thinking-and-working-politically-doing-development-differently-etc-getting-beyond-the-peoples-front-of-judea/).

Green, D. and Guijt, I. (2019) *Adaptive programming in fragile, conflict and violence-affected settings. What works and under what conditions? The case of Institutions for Inclusive Development, Tanzania*. Case Study for the Action for Empowerment and Accountability Programme. Sussex: Itad and Oxfam in association with the Institute for Development Studies (IDS) (https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/14562).

Hudson, D. and Marquette, H. (2015) 'Mind the gaps: what's missing in political economy analysis and why it matters' in A. Whaites (ed.) *A governance practitioner's notebook*. Paris: Organisation for Economic Co-operation and Development (OECD) (www.oecd.org/dac/accountable-effective-institutions/Governance%20Notebook%201.3%20Hudson%20et%20al.pdf).

I4ID – Institutions for Inclusive Development (n.d.) 'Education systems for the deaf'. Briefing note. (www.i4idtz.org/wwwo-inced).

Kelsall, T. (2018) *Thinking and working with political settlements: the case of Tanzania*. London: ODI (www.odi.org/publications/11234-thinking-and-working-political-settlements-case-tanzania).

Kirsch, R., Siehl, E. and Stockmayer, A. (eds) (2017) *Transformation, politics and implementation: smart implementation in governance programs*. Frankfurt: Nomos/GIZ.

Kitson, N. (2019) *Informal payments in the public health sector in Guinea-Bissau*. London: ODI (www.odi.org/publications/11464-informal-payments-public-health-sector-guinea-bissau).

Laric, M. and Waddell, N. (2016) 'Nepal: hydropower and the politics of investment' in D. Booth (ed.) *Politically smart support to economic development: DFID experiences*. London: ODI (www.odi.org/publications/10357-politically-smart-economic-development-nigeria-nepal).

Laws, E. (2018) *Thinking and working politically in Somalia: a case study on the Somalia Stability Fund*. ODI Report. London: ODI (www.odi.org/publications/11136-thinking-and-working-politically-somalia-case-study-somalia-stability-fund).

Laws, E. (2020a) *Working effectively through partnerships: lessons from Institutions for Inclusive Development in Tanzania*. London: ODI (www.odi.org/publications/16730-lessons-institutions-inclusive-development-tanzania).

Laws, E. (2020b) *Rethinking value for money for adaptive, politically smart programmes: lessons from Institutions for Inclusive Development*. London: ODI (www.odi.org/publications/16730-lessons-institutions-inclusive-development-tanzania).

Laws, E. (2020c) *Mainstreaming gender in an adaptive, politically smart governance programme: lessons from Institutions for Inclusive Development in Tanzania*. London: ODI (www.odi.org/publications/16730-lessons-institutions-inclusive-development-tanzania).

Laws, E. and Marquette, H. (2018) *Thinking and working politically: reviewing the evidence on the integration of politics into development practice over the past decade*. Birmingham: University of Birmingham Thinking and Working Politically Community of Practice (https://twpcommunity.org/wp-content/uploads/2018/04/Thinking-and-working-politically-reviewing-the-evidence.pdf).

Laws, E. and Valters, C. (2021) *Value for money and adaptive programming: approaches, measures and management*. ODI Working Paper. London: ODI (www.odi.org/publications/17914-value-money-and-adaptive-programming-approaches-measures-and-management).

Levy, B. (2014) *Working with the grain: integrating governance and growth in development strategies*. Oxford: Oxford University Press.

Levy, B. and Walton, M. (2014) *Institutions, incentives and service provision: bringing politics back in*. ESID Working Paper No. 18. Manchester: Effective States and Inclusive Development Research Centre (https://assets.publishing.service.gov.uk/media/57a08a47e5274a31e0000506/esid_wp_18_levy-walton.pdf).

Lindblom, C.E. (1959) 'The science of "muddling through"' *Public Administration Review* 19(2): 79–88 (https://doi.org/10.2307/973677).

Lindell, I., Norström, J. and Byerley, A. (2016) 'New city visions and the politics of redevelopment in Dar es Salaam'. Working Paper, Nordiska Afrikainstitutet (NAI). Uppsala: NAI (www.diva-portal.org/smash/record.jsf?pid=diva2%3A1070397&dswid=9045).

Marquette, H., Hudson, D. and Waldock, N. (2016) *Everyday political analysis*. Birmingham: Developmental Leadership Program (DLP), University of Birmingham (www.dlprog.org/publications/research-papers/everyday-political-analysis).

MOEST – Ministry of Education, Science and Technology (2017) 'National strategy for inclusive education 2018–2021'. The United Republic of Tanzania (www.rodra.co.za/images/countries/tanzania/policies/National%20Strategy%20for%20Inclusive%20Education_version%20for%20signing.pdf).

Piron, L.H., Baker, A., Savage, L. and Wiseman, K. (2016) *Is DFID getting real about politics? A stock take of how DFID has adopted a politically-informed approach (2010–2015)*. London: Department for International Development.

Ramalingam, B., Jones, H., Reba, T. and Young, J. (2008) *Exploring the science of complexity: ideas and implications for development and humanitarian efforts*. London: ODI (www.odi.org/publications/583-exploring-science-complexity-ideas-and-implications-development-and-humanitarian-efforts).

Ramalingam, B., Laric, M., and Primrose, J. (2014) *From best practice to best fit. Understanding and navigating wicked problems in international development*. London: ODI (www.odi.org/publications/8571-best-practice-best-fit-understanding-and-navigating-wicked-problems-international-development).

Ramalingam, B., Wild, L. and Buffardi, A. (2019) *Making adaptive rigour work: principles and practices for strengthening monitoring, evaluation and learning for adaptive management*. London: ODI. (www.odi.org/publications/11311-making-adaptive-rigour-work-principles-and-practices-strengthening-mel-adaptive-management).

Rocha Menocal, A. (2014) *Getting real about politics: from thinking politically to working differently*. London: ODI (www.odi.org/publications/8325-getting-real-about-politics-thinking-politically-working-differently).

Rocha Menocal, A. (2020) *Why does inclusion matter? Assessing the links between inclusive processes and inclusive outcomes*. London: ODI (www.odi.org/publications/16782-why-does-inclusion-matter-assessing-links-between-inclusive-processes-and-inclusive-outcomes).

Teskey, G. and Tyrrel, L. (2017) *Thinking and working politically in large, multi-sector facilities: lessons to date*. Governance Working Paper Series Issue 2. ABT Associates (https://abtassocgovernancesoapbox.files.wordpress.com/2017/11/abt-associates-governance-working-paper-series-issue-no-2-final-171120.pdf).

Unsworth, S. (2015) 'It's the politics! Can donors rise to the challenge?' in A. Whaites, E. Gonzalez, S. Fyson and G. Teskey (eds) *A governance practitioner's notebook: alternative ideas and approaches*. Paris: OECD (www.oecd.org/dac/accountable-effective-institutions/Governance%20Notebook.pdf).

Wild, L., Booth, D., Cummings, C. et al. (2015) *Adapting development: improving services to the poor*. London: ODI (www.odi.org/publications/8125-adapting-development-improving-services-poor).

Wild, L., Booth, D. and Valters, C. (2017) *Putting theory into practice: how DFID is doing development differently*. London: ODI (www.odi.org/publications/10729-putting-theory-practice-how-dfid-doing-development-differently).

Wild, L. and Foresti, M. (2011) *Politics into practice: a dialogue on governance strategies and action in international development*. London: ODI (www.odi.org/publications/5860-politics-practice-dialogue-performance-strategies-and-action-international-development).

ODI is an independent, global think tank, working for a sustainable and peaceful world in which every person thrives. We harness the power of evidence and ideas through research and partnership to confront challenges, develop solutions, and create change.

ODI
203 Blackfriars Road
London SE1 8NJ

+44 (0)20 7922 0300
info@odi.org

odi.org
odi.org/facebook
odi.org/twitter