Water Resour Manage
DOI 10.1007/s11269-015-1081-0

Scaling up of Policy Experiments and Pilots: A Qualitative Comparative Analysis and Lessons for the Water Sector

Sreeja Nair[1] · **Michael Howlett**[1]

Received: 19 December 2014 / Accepted: 28 July 2015

Abstract The use of experimentation by practitioners and resource managers as a policy instrument for effective policy design under complex and dynamic conditions has been well-acknowledged both in theory and practice. For issues such as water resource management, policy experimentation, especially pilot projects, can play an important role in exploring alternate courses of action when faced with long-term uncertainty. While the political aspects of experimentation design and outcomes have been alluded to by several policy scholars, there is lack of empirical evidence that explores their interplay with other factors that may also be critical for scaling up of policy experiments. This paper examines experiences with scaling up of different types of water policy experiments through a Qualitative Comparative Analysis of fifteen pilot initiatives in multiple sectors. Presence of political support is found to be necessary for scaling up in 97 % of the cases studied, followed closely by the need for synergies with ongoing policies and programmes. When in combination with effective pilot planning and strong monitoring and evaluation, both these factors create a sufficient condition for successful scaling up in nearly 60 % of the cases studied.

Keywords Policy experiments · Policy pilots · Scaling up · Water policy · Water management · Policy formulation

1 Introduction

Effectively managing water resources is a major challenge for policymakers and water managers given the multiple stresses which are adversely impacting water resources worldwide. These impacts are manifested in the form of conflicts over water allocation and use, inadequacy of current water distribution systems, the presence of multiple stakeholders with

✉ Michael Howlett
howlett@sfu.ca

[1] LKY School of Public Policy, National University of Singapore, Singapore, Singapore

Published online: 03 September 2015

varied interests which lead to competing demands for resource use, and increasing numbers and severity of environmental stressors such as climate change which impact on the water cycle (Moore et al. 2014).

In addition to current challenges, policymakers and resource managers also need to consider how the impacts of current and likely new stressors will affect water resources over longer time horizons in order to undertake effective anticipatory policy planning. Policy experiments form a useful policy tool to manage such complex long-term policy issues by aiding the ex-ante evaluation of policies, generating learning outcomes and policy relevant information under dynamic conditions (McFadgen 2013). Enhanced experimentation and consequent learning can also aid in adapting to the "dynamic drivers and expressions of risk" in a changing policy environment (O'Brien et al. 2012).

Pilot projects in particular are an increasingly common mode of policy experimentation and a widely used method to introduce major government policies or programmes in a phased manner, allowing them to be "tested, evaluated and adjusted" beforehand (Cabinet Cabinet 2003). However there exist several challenges in translating or 'scaling-up' of experimental projects, including pilots, which are not well understood in the water or any other sector (Stoker 2010).

One important challenge in terms of their ability to act as a predictive method beyond the context in which these are applied involves difficulties encountered in 'scaling-up' or moving from the pilot to full project or programme status. Spicer et al. (2014) argue "scaling up is a craft not a science" alluding to the predominant political nature of the activity compared to its technical aspects, but this remains merely an assertion and the relevance of political versus other factors is not well known or understood in this process.

Several experiments have been undertaken in the water sector at the national and local scales which are able to shed some light onto and provide critical insights into the factors which affect the scaling-up process. This paper presents an analysis of these cases and the factors that can influence scaling up of policy experiments, including pilots and draws lessons for experimentation. This is done through a review of selected examples of water policy experiments and a Qualitative Comparative Analysis of several pilots.

1.1 What is a Policy 'experiment'?

The design of policy experiments is not only a technical process, but highly driven by the interests, behavior and attitudes of the stakeholders. Compared to the earlier works however that focused more on the *content* of the experiments itself, the more recent literature on experimentation has shifted its focus to the *process* of experimental policy design, including the role of various stakeholders therein (van der Heijden 2013). This new wave of "experimentalist governance" presents an iterative process of "provisional goal setting" with the intention of revising the goals based on the learning derived from trying out alternate modes of goal achievement in different contexts (Sabel and Zeitlin 2012).

Results from laboratory and field experiments have often been used by scholars to provide policymakers and practitioners with evidence of the impacts of selected experimental interventions as well of their feasibility and acceptability by key stakeholders, including their intended beneficiaries. The broad classification of experimental projects by Rondinelli (1993) helps characterize different types of water policy experiments and is used to guide the discussion in this paper.

Water policy experiments that fall into the first category of experimental projects based on Rondinelli's classification include activities such as *need-based assessments*. A common form

of such assessments has been willingness- to-pay surveys which are often used as a proxy to assess the demand for services such as water and sanitation. Pattanayak et al. (2006), for example, conducted a willingness-to -pay experiment of 1800 households in Sri Lanka to demonstrate that demand for improvement in water and sanitation services is driven by a combination of several factors such as socio-economic status, costs, location, means for self-provision and perceptions of stakeholders. Results of the experiment also indicated that while presence of policy incentives such as connection fee subsidies can increase the demand for piped water, the question of whether the benefits accrued by scaling up are more than the costs incurred still remained. Behavioural variables at the level of the individual can form key decisive factors in influencing the overall outcome of this category of policy experiments. While behaviour can be regulated with incentives to some extent, there are limitations to how observations at the local level can be considered to be a good indicator for the overall success or failure of the experiment when it is scaled up.

The second category of policy experiments involves projects that explore the most effective way of achieving pre-set policy goals. Under dynamic conditions this can involve *innovations and transition experiments*. In the past decade the field of transitions management has gained prominence to explore "a range of possible pathways for change" (Farrelly and Brown 2011).

Transitions can be defined as 'a gradual, continuous process of structural change within a society or culture' and are complex, spread over long timeframes, involve multiple actors and occur across multiple levels (Rotmans et al. 2001). Transitions require "steering, facilitation and coordination" and experimentation and learning form important concepts (Farrelly and Brown 2011). In the context of urban water sustainability Farrelly and Brown (2011) examined eleven local-scale experiments in Australian cities and found sustainable transitions to urban water management required changes in underlying culture and beliefs along with structural reforms. The role of 'bridging organizations' was found to be critical to collate insights from local-scale experiments and inform future policy and practice. In another study from Hyderabad city in India, Nastar (2014) explores the impacts of ongoing legislative, technical, managerial, and social aspects of the urban water regime on the citizens' access to water. The study finds that scaling-up of innovative 'niche experiments' that aim towards transitions in urban water management is often impeded due to system lock-ins and tendency of donor agencies as well as current water policy and urban development initiatives to preserve status-quo.

The third category of policy experiments involves those that aim at *identification of gaps* in current policy practices. While pilots form a useful means to investigate gaps, this intended purpose is not met if errors or gaps identified in the pilot phase are not corrected before scaling up of these pilots. For example, privatization of urban water services provision in Kenya began on an experimental basis followed by large-scale expansion. The objective of these privatization efforts were to decentralize water governance structure to alleviate problems such as unaccounted for water losses, unmetered water usage and uneconomic water usage fee and rates that were linked to a highly centralized water governance. However, the experiment as well as scaling up has not achieved its intended outcomes because both at the local and city scale the privatization efforts were unable to avoid intrusion of central and local government authorities in its functioning. A second cycle of privatization was also attempted but errors/ gaps identified in the first phase were not considered in subsequent efforts, thereby rendering the experiments futile (Akumu and Appida 2006).

A fourth category of policy experiments is *natural experiments*. While studying from history i.e., natural experiments in the water sector is helpful however their applicability as a 'blueprint' for similar outcomes in the future is limited for dealing with policy issues such as climate change that face a high degree of uncertainty. The key challenge is that under 'surprise' (Walker et al. 2010; Lempert et al. 2003) these experiments offer little or no scope for decision-makers to respond from history or experience.

The final type of policy experiments Rondinelli identified focusses on problems that are partly or wholly undefined. Typically under such conditions of uncertainty, *policy pilots* are undertaken. The term policy pilots as used in this paper refers to projects initiated by governments for policy purposes, including testing potential policies, implementing policies that have difficulties in being implemented and evaluating new policies at an early stage. This is the principle type of experiment undertaken in the water sector and is the subject of the QCA analysis which follows. The next section discusses the characteristics of the scaling up process of policy pilots in general as a mode of policy experimentation.

1.2 Scaling up of Policy Pilots: Factors and Barriers

Hartmann and Linn (2007) define scaling up as "expanding, replicating, adapting and sustaining successful policies, programs or projects in geographic space and over time to reach a greater number of people". Scaling up occurs when a program increases in size, its geographical spread or budget (quantitative); increases in its range of activities and interaction with related programs (functional); increases in political power and engagement with wider political processes (political) or increases in organizational capacities and processes (organizational) (Gillespie 2004).

Empirical evidence on the composition of effective policy pilots and the process of their diffusion – that is, their continuation or expansion - is generally lacking (Vreugdenhil et al. 2009). Many factors are thought to influence pilot dynamics, however, including the pilot design and the context (Vreugdenhil, 2010). These can include factors such as the number and type of stakeholders involved that further influences the availability of knowledge and resources, the choice of scale of implementation and the choice for pilot sites, the mode of governance that influences the nature of stakeholder engagement and learning, the level of innovativeness of the pilot and how it converges or diverges from the current policy context, the degree of flexibility present to make changes to adapt to local conditions and finally the timing of the strategy for pilot diffusion. Scaling up of pilots and their sustenance beyond pilot sites operates often operates in conjunction with, and needs, sustained efforts towards empowerment and capacity building of local communities and beneficiaries of the pilot (Turton and Bottrall 1997). Pilot diffusion can face impediments when a strategy for diffusion management is entirely absent or poor, or when there is widespread opposition from some critical stakeholders (Vreugdenhil 2010). Additionally, if the policy change involves significant costs it is likely to motivate policymakers to resist change and thus increase the 'stickiness' of existing policies (Callander 2011).

Scaling up, whether in space or time, also often runs the risk that the initial project objectives and outcomes become less appropriate or relevant for the new context (Simmons et al. 2007). For example, Margerum (2012) presents the successful case of watershed management at the state level in Oregon, United States and argues that success at the watershed level may be rather fragmented and thus may not uniformly translate or scale-up to a larger i.e., river basin scale owing to limitations in stakeholder capacities and quality of

coordination efforts. Furthermore, successful small-scale, often non-regulatory approaches such as water management efforts taken at a watershed level might not always be scaled-up successfully to address issues at the larger scale such as river-basin flooding.

2 Framework for the Study

As Hartmann and Linn (2007) suggest, the key challenge in studying pilot projects and their outcomes is to identify both context-specific as well as universal elements contributing to scaling up and to ensure general elements are maintained while leaving scope for context-specific changes to take shape through adaptation and learning (Hartmann and Linn 2007).

Based on the study of barriers and key factors set out above, they also identify seven elements as being critical for scaling up of developmental interventions. These factors are (i) applying leadership, vision and values; (ii) managing political constituencies; (iii) ensuring supportive policies; (iv) developing institutional capacity; (v) creating incentives and accountability; (vi) practicing evaluation, learning and feedback; and (vii) planning for success (see Table 1).

Table 1 Key factors for scaling up (Hartmann and Linn 2007)

Factors	Description
Leadership, vision and values	Presence of leaders driving the scaling up with a clear vision, enabling institutions to exemplify a set of values for achieving scaling up, to avoid "short-termism" of programs and "fragmentation of effort".
Presence of political constituencies	Scaling up is supported by political constituencies. This entails the active engagement of political players in the scaling up process and its placement on their agendas, driven by "need and appropriateness" rather than any personal interest and guarded from "elite capture".
Presence of supportive policies, programs and projects	Presence of a supportive policy framework (laws, regulations, norms and linkages with related policies, programs and projects) for scaling up.
Strong institutional support and capacities to facilitate change	Scaling up requires adequate institutional and human capacities and additional training and development and institutional capacity building. These efforts also need to be constantly evaluated in their performance relative to appropriate benchmarks, while ensuring accountability.
Incentives and accountability	Incentives for stakeholders form a critical factor for enabling leadership, political support and institutional capacity for scaling up. Accountability for scaling up on the other hand is essential to ensure that incentives are in sync with some shared objectives of the stakeholders.
Effective monitoring & evaluation (M & E)	Monitoring and Evaluation focusing on scaling up as a key indicator of success can assess the impact of the program and obtain feedback for improvement, and thus build a case for garnering political and stakeholder support and sustainability of the program.
Scaling up benefits from an orderly and gradual process of planning	A systematic and gradual process, careful planning, and clear demarcation of roles and responsibilities of partners and strong communication channels are important factors for scaling up.

This framework is applied to selected cases of policy pilots in the water sector to further investigate the characteristics of factors that can influence scaling up of these pilots, especially their group or synergistic effects. A Qualitative Comparative Analysis (QCA) is presented towards this end. Using a fuzzy-set Qualitative Comparative Analysis this paper identifies sets of factors that are conditional and sufficient for the scaling up of the selected policy pilots.

3 A Qualitative Comparative Analysis of Selected Policy Pilots

3.1 Narrative Review

A narrative review "summarizes different primary studies from which conclusions may be drawn into a holistic interpretation contributed by the reviewer's own experience, existing theories and models" (Kirkevold 1997). Results are of a qualitative meaning and help synthesize the "diversities and pluralities of understanding around scholarly research topics" (Jones 2004).

A narrative review is undertaken for this paper because there is lack of theoretical frameworks and synthesis of evidences from the large number of operational/ abandoned pilots to glean common factors that make pilots scale up. This review of published articles on pilot projects in different countries and sectors was conducted using Googlescholar. A long time period (2000–2015) was selected with the aim of obtaining more documentation on pilots as these are rarely reported, especially in cases when they are not successful.

A combination of the following keywords was used to conduct the online search for articles: 'water policy experiments', 'scaling up', 'policy pilots', 'success', 'failure', 'diffusion of policy pilots', 'evaluating policy pilots', 'replication'. Articles on specific pilots were identified after a review of the abstract that were selected based on the following criteria:

- Articles should be first-hand documentations of individual pilot projects with detailed analysis
- Articles should refer to pilots that were consciously launched with the objective of scaling up and guiding future policy development.

The analysis has been extended beyond the water sector with the aim of identifying certain common factors that influence scaling-up of pilots. A total of fifteen cases were identified of which seven pilots span health care, poverty alleviation and agriculture risk management, with different levels of diffusion and designed and governed in different ways and eight related to the water sector.

3.2 Fuzzy-set Qualitative Comparative Analysis

Qualitative Comparative Analysis (QCA) is an analytical technique based on Boolean algebra to allow for comparison of qualitative cases that are often large enough to do in-depth qualitative analysis but too small to do variable-oriented quantitative analysis.[1] QCA has been

[1] For further details see http://www.u.arizona.edu/~cragin/fsQCA/index.shtml.

used for studies in the water sector. For example Aubin and Varone (2013) use QCA to study the conditions under which new water users obtain access to the water resource and consequent changes in the behavior and water use practices of existing users. Issues of institutional arrangements related to water resource management have also been explored using QCA. For example, Kosamu (2014) uses QCA to study factors that lead to sustainability of community managed small-scale fisheries in a wetland in Southern Malawi, in the absence of any state involvement.

The objective of QCA is to enable causal interpretation in addition to detailed qualitative information that is obtained from case studies, in order to understand the different combination of plausible factors that could lead to a specific outcome (Ragin 2008).

QCA is particularly helpful in instances where there is an hypothesis regarding the underlying causal factors affecting the outcome being studied (scaling up in this case), when different combinations of these plausible causal factors could give rise to the outcome and conditions are sufficient only when they are in combination, when results need to be interpreted as "necessary and sufficient conditions", when the number of cases is very low for conventional quantitative methods to be applied, when a good deal is known regarding the cases, and when the key concepts are clearly defined and measured (Ragin 2008). The variables in QCA are either presented as "crisp" i.e., binary sets that denote presence or absence (1 or 0 respectively) of "membership" in a specific category or "fuzzy" sets which split this all-or-none categorization into further categories using scores from 0 to 1 (Ragin 2006). This paper uses fsQCA (software fsQCA 2.0).[2]

4 Methods

The steps followed in the fsQCA for this paper are given below:

4.1 Constructing a Data Matrix

The first step in fsQCA is to construct a data matrix, which lists certain characteristics of the cases as variables. These characteristics denote the "degrees of membership" of a defined category. Fuzzy set can allow for scores in between 0 and 1 to denote various degrees of membership. For this study a 4-point fuzzy set has been used where the membership in a particular category has been denoted in the following way:

$$1 = \text{fully in},\ 0.67 = \text{more in than out},\ 0.33 = \text{more out than in},\ 0 = \text{fully out} \quad (1)$$

Here the data matrix consists of the fifteen case studies as rows and the seven hypothesized causal conditions (as per the framework discussed in section 2) factors to be tested as columns, including an additional column called 'scaling up'. This column marks whether scaling up has happened or not and to what extent based on categorization discussed in (1).

[2] Free to download at http://www.u.arizona.edu/~cragin/fsQCA/software.shtml

4.2 Constructing the Truth Table

Next, a "truth table" is constructed marking 'scaling up' as the 'outcome' that the paper aims to study based on membership scores of causal 'conditions' that may be necessary or sufficient for the outcome to happen. The truth table considers each case as a combination of the characteristics selected. Normally, four kinds of result can be expected in the truth table:

- A Combination of specific characteristics lead to positive outcomes,
- A Combination of specific characteristics lead to negative outcomes,
- There are contradictory cases i.e., a specific combination leads to positive outcomes in some cases and negative in others, and
- There are no cases for specific combinations: This is likely for small-n studies, wherein there will be many combinations of characteristics that are possible but not observed in any of the cases (due to the small sample size). Hence in these cases it is also not possible to say whether the outcome occurred or not (termed 'remainders' in fsQCA).

This study is a small-n type hence the remainders are excluded from the analysis. There are also no contradictory cases that were found. Studying the truth table in this case thus can give a big picture of the variety of combinations of characteristics that are common or those that happen often or seldom.

4.2.1 Coding of Outcome and Causal Conditions

Determination of crossover points to code the causal conditions and outcome is challenging (van der Heijden 2013). For purposes of this paper scores are assigned to the seven hypothesized causal conditions and the observed outcomes in terms of their scaling-up, based on the scores provided in Table 2. These scores were assigned to compare and characterize these pilots relative to each other.

The values for the fifteen cases are assigned after reviewing the cases and understanding the case context. Values are assigned for the degree of presence of each of the seven hypothesized causal conditions for scaling-up. Scaling-up (outcome) refers to quantitative scale-up in this paper i.e., an increase in geographic spread of the pilot and coverage of intended beneficiaries.

Table 3 presents an overview of the fifteen cases that have been considered in this paper.

4.3 Analysis of Sufficient and Necessary Conditions

The 'truth table' presents the different combinations of causal factors that have met specified criteria of sufficiency for the outcome to occur. This suggests that the membership score on the outcome is always higher than the membership score of the causal combination. The analysis of necessary conditions in fsQCA assesses individual causal factors that may be necessary for the outcome to occur. This suggests that the membership score on the outcome is usually always lower than the membership score of the causal factor being investigated. In other words, when X (causal factor) is considered as a necessary condition for Y (outcome) to occur, it means that Y cannot occur without X, i.e., Y (outcome) is a subset of X (causal factor). On the other hand, when X is considered as a sufficient condition means that if Y is present X must be present too. This however does not mean that X by itself will cause Y (i.e., there may be

Table 2 Coding of outcome and causal conditions

Fuzzy score	Outcome=Scaleup	Leadvision	Polsupp	policies	instcap	incentives	moneval	Planning
0	No scaling up	Role of specific leaders is not explicitly documented	Resistance from major political players to scaling-up	No supportive policies	No institutional support and training or capacity building initiatives	No incentive structure or accountability mechanism	No M&E framework present	No prior plan for scaling-up
0.33	Low geographic spread and low coverage of intended beneficiaries	Leaders present only in initial phases	Some major political players and stakeholders support while others resist scaling up	Very few policies available in support	Little emphasis on institutional strengthening, capacity building and training	Weak incentive structures and accountability mechanisms	Weak M& E	Vague planning, demarcation of roles and responsibilities, poor communication strategy
0.67	Complete scale-up in terms of geographic spread but low coverage of intended beneficiaries	Leaders present but keep changing	Moderate support from major political players and stakeholders	Moderate synergies with existing policies; efforts being made to increase this support	Moderate emphasis on institutional strengthening, capacity building and training	Moderate incentive structures and accountability mechanisms	Good M& E in place but with constant revisions	Good planning, demarcation of roles and responsibilities and communication strategy, but with constant revisions
1	Complete scale-up in terms of geographic spread and coverage of intended beneficiaries	Strong, stable leadership from the beginning to end of the pilot	Strong support from major political players and stakeholders	High convergence with ongoing policies, pilot is receiving support from these policies	Major focus on institutional strengthening, capacity building and training	Strong incentive structures and accountability mechanisms	Strong and regular M& E	Clear planning, demarcation of roles and responsibilities, and communication strategy

Table 3 Overview of cases

S. no.	Country and Pilot	Objective	Outcome= Scaling up	Score
1	China- Methadone Maintenance treatment (MMT)	To increase the coverage of MMT, its beneficiaries and improve accessibility of services	The project moved from a pilot in 8 sites in 2004 to a nation-wide programme covering 27 provinces by the end of 2009.	0.67
2	India- Kudumbashree	A multi-sectoral poverty alleviation program initiated by the Government of Kerala (GoK), India to eradicate poverty in the state by 2008.	In 1991, the GoK, India and UNICEF initiated Community-Based 1 Nutrition Program (CBNP) in Alleppey town to improve the health and nutritional status of children and women. In 1998, GoK scaled up the program to the entire state in 1998.	1
3	Vietnam- Injectable contraception and quality of care	To scale-up introduction of the injectable contraceptive depot-medroxyprogesterone acetate (DMPA) as part of health intervention packages to improve the quality of care in the family planning programme.	After pilot testing of the interventions in three provinces of Vietnam in 1996, these were scaled up to 21 and then all of the 64 provinces by 2002.	1
4	Cambodia- Health Equity Funds (HEFs)	To use HEFs for translation into health policies for the poor to promote equity	HEFs pilots were initiated in 2000 in two urban slums and were translated into a national health policy; and scaled up to 50 HEF schemes based in 51 hospitals in Cambodia.	1
5	China- Family planning (FP)	To promote family planning and limit births as part of China's sustainable development goals To	Initiated in 6 counties of China in 1995 and scaled into a national reform effort by 2004	1
6	Ghana- Community-based Health Planning & Services (CHPS)	guide national reforms for supporting community-based primary health care	Pilot launched in 1994 in three villages of Ghana. By 2003 the pilot became a national initiative for district planning process and community-based health care.	0.67
7	Thailand- 100% condom programme	To control Sexually-Transmitted Diseases in sex workers	Initiated in one province in 1989 and implemented in all provinces in 1991 owing to the success of the pilot.	1
8	Bangladesh-Community Led Total Sanitation	To bring about behavioral change to initiate community-led total sanitation (CLTS)	CLTS developed from a pilot in few districts (1999) to a national program on sanitation (2003) that embraced many principles of CLTS.	0.67

Table 3 (continued)

S. no.	Country and Pilot	Objective	Outcome= Scaling up	Score
9	India- *Swajaldhara*	To create a demand-driven sector reform to provide safe drinking water to all citizens, with the community contributing 10% to installation costs and looking after subsequent operation and maintenance costs	Sector-reform pilot (SRP) project initiated in 1999 moved into a countrywide programme, *Swajaldhara* in 2002.	0.67
10	Zambia: Urban poverty alleviation-Programme of Support for Poverty Elimination and Community Transformation (PROSPECT)	To alleviate urban poverty through empowerment of poor communities and enabling their participation in decision making and building collective capacities to act.	The PROSPECT ended in 2004 as a 6-year follow-up to two previous projects that operated one after the other from 1992. PROSPECT operated in only 13 of Lusaka's total 37 compounds, reaching 300,000-400,000 of the estimated population of 800,000 in Lusaka's informal settlements.	0
11	Pakistan- The Orangi Pilot Project (OPP)	To provide sanitation for the urban poor in Orangi, one of Karachi's poorest districts with the largest informal settlement population of 1.2 million.	By April 2001 OPP had covered almost 90% of the entire settlement, since its inception in 1980.	0.67
12	Tanzania- Payment for ecosystem services (PES)	To demonstrate a watershed payment in Uluguru mountains for ecosystem services scheme that links rural ecosystem service providers to urban water users through economic transfers to help maintain water supply and forest habitat.	Though the project has been successful in putting working arrangements in place and initiating the first pilot water payments in eastern Africa, significant scaling up is needed if measurable changes in water quality are to be achieved across the pilot site.	0.33
13	South Africa- Saldanha Bay Water Quality Forum Trust pilot	To implement existing national policy and to address local coastal management issues for stimulating innovation	This pilot is widely promoted in national best practice guides as a model for local institutional arrangements, e.g., in 'South Africa's operational policy of the disposal of land-derived wastewater to the marine environment'	0.67
14	Australia- Urban water planning	To develop and implement a multi-disciplinary participative approach to urban water planning in the Cooks River Catchment	The pilot operated from 2002-2011 and helped in uniting municipalities in active support for sustainable water practices at	0.67

Table 3 (continued)

S. no.	Country and Pilot	Objective	Outcome= Scaling up	Score
			a political level, replicated in other areas. Its underlying values and practices are being institutionalized through new governance rules and structures within the catchment.	
15	Denmark-Decentralized drainage	To implement decentralized drainage solutions	Moved from a pilot in Egedal municipality to a national level sewage reform plan	1

other factors influencing Y too). In other words, in this case X (causal factor) becomes a subset of Y (outcome) (Ragin 2008).

4.4 Analysis of Consistency and Coverage

Some other results that are provided by the QCA include Consistency i.e., the degree to which the cases sharing a specific combination of causal factors share the same outcome and coverage, i.e., the degree to which a specific causal combination accounts for occurrence of an outcome. Raw coverage measures the "proportion of memberships in the outcome explained by each term of the solution". Unique coverage measures "the proportion of memberships in the outcome explained solely by each individual solution term (i.e., memberships that are not covered by other solution terms)" (Ragin 2006).

5 Findings

The results following the fsQCA and in-depth analysis of the cases are presented in this section in three parts:

(1) Results of the truth table which is constructed by putting the values of occurrence of the outcome i.e., scaling up and degree of presence of the seven potential causal factors that are being tested.
(2) Results from assessment of the necessary conditions
(3) Results from assessment of the sufficiency conditions

5.1 Assessment of the Truth Table

The truth table is presented in Table 4. It shows the cases on the left-hand side. The next column shows the outcome, i.e., in this case the degree to which scale up has

Table 4 Truth table

Case	Outcome= Scaleup	leadvision	Polsupp	policies	instcap	incentives	moneval	Planning
China MMT	0.67	0	1	1	0.67	0	1	0.67
India Kudumbashree	1	0.67	1	1	0.67	0.33	0.67	0.33
Vietnam DMPA	1	0	1	0.67	0.67	0	0.33	0.67
Cambodia HEF	1	0	1	1	0	0	1	1
China family planning	1	0.67	1	0.67	0.67	0	1	0.67
Ghana CHPS	0.67	0.33	0.67	1	0.67	0	1	0.67
Thailand condom programme	1	0	1	1	0.67	0	1	0.67
Bangladesh CLTS	0.67	1	0.67	0.33	0.67	1	0.33	0.33
India Swajaldhara	0.67	0.33	0.67	0.67	0.33	1	0.67	0.33
Zambia PROSPECT	0.33	0.33	0.33	0.33	0.33	0	0.33	0.33
Pakistan OPP	0.67	0.67	0.67	0.33	0.67	0	0.33	0.33
Tanzania PES	0.33	0.33	0.33	0.67	0.33	1	0.67	0.67
South Africa- Saldanha bay	0.67	0.67	1	1	0.67	0.33	0.33	0.33
Australia Urban water planning	1	1	0.67	0.67	0.67	0.67	0.33	0.67
Denmark decentralized drainage	1	0.67	1	0.67	0.67	0.33	0.33	0.33

Legend: *scaleup* outcome i.e., degree of scale-up, *leadvision* presence of leader, vision, values, *Polsupp* political support, *policies* synergy with current policies/ programs, *instcap* institutional support and capacities; *incentives* presence of incentives and accountability, *moneval* monitoring and evaluation, *planning* gradual process with detailed planning, clear communication and engagement of stakeholders with clear demarcation of roles and responsibilities

occurred (or not) in the particular case. These scores for the 7 hypothesized causal conditions are used to indicate the degree to which the causal condition was present in that case.

5.2 Assessment of Necessary Conditions

The fsQCA is used to run an assessment of necessary conditions. The results are presented in Table 5. A score of above 0.8 is considered to be good for acceptance of a causal factor as a necessary condition (Kent 2008).

Table 5 can thus be interpreted in the following way. Looking at the consistency values, political support is necessary for scaling up in 97 % of the cases studied. This is closely followed by policy synergies in 83 % of the cases, and institutional capacities and M&E which is necessary for scaling up in 71 % of the cases studied.

Table 5 Assessment of necessary conditions

Factor	Consistency	Coverage
leadvision	0.54	0.95
polsupp	0.97	0.94
policies	0.83	0.88
Instcap	0.71	1
incentives	0.29	0.71
moneval	0.71	0.89
planning	0.65	0.96

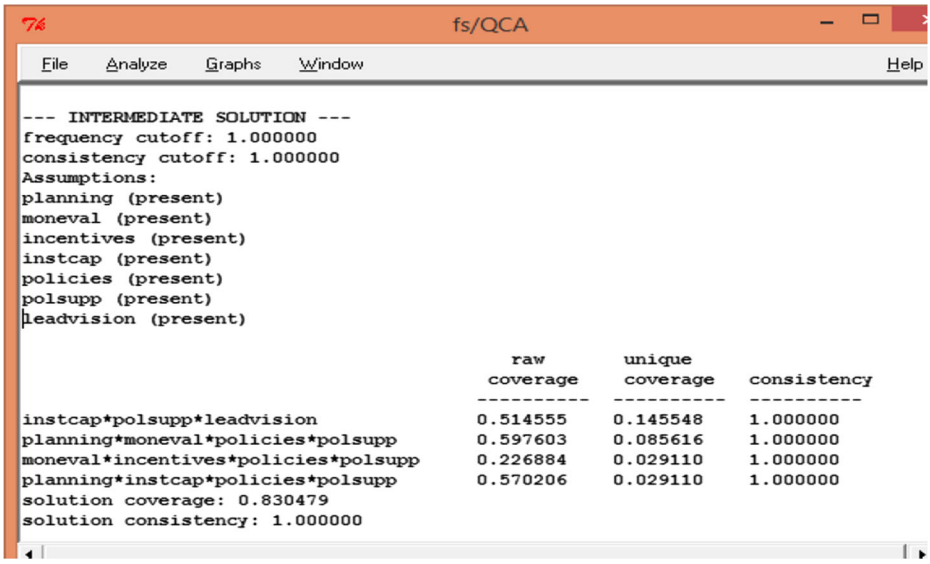
5.3 Assessment of Sufficiency Conditions

The truth table is now analyzed using the fsQCA software for sufficient conditions (Table 6). Based on the hypothesis for this paper, the model used suggests that scale-up is considered to be a factor of all 7 causal conditions i.e.,

Scaleup=f (planning, moneval, incentives, instcap, policies, polsupp, leadvision)

Results presented in Table 6 can be interpreted as follows:
Firstly, presence of either of the following combinations
(instcap*polsupp*leadvision)
(planning*moneval *policies*polsupp)
(moneval*incentives*policies*polsupp)
(planning*instcap*policies*polsupp)
is completely consistent i.e., are *sufficient* to ensure scale-up.

Table 6 Assessment of sufficient conditions



Secondly, the *raw coverage* value indicates that the presence of (planning*moneval *policies*polsupp) can explain 59.7 % of the scaling-up that occurs. The *Unique coverage* value indicates that when only (instcap*polsupp*leadvision) is present it can explain 14.5 % of the occurrence of scale-up in the cases studied.

6 Conclusions

Policy experimentation has been well-acknowledged as a policy tool used to deal with complex and dynamic policy issues. Different types of policy experiments including pilot projects have been conducted in the water sector and these have provided useful insights to water resource managers for policy design.

Policy pilots may be scaled up in space, time or based on their purpose. Individual factors that are considered important for scaling up can be studied in combination with others to see the impact various combinations can have on scaling up. This can be particularly important when governments and other agencies operate with limited resources, and thus can invest these resources in enhancing these specific factors in a targeted manner. But these processes and the factors which drive them are not well understood.

Pilots might sometimes be used as an excuse to garner political acceptability, or maybe abandoned citing them as failures because the political milieu might not be conducive for it to move ahead. Presence of multiple stakeholders and their power positions can also influence the scaling up process. Many water policy experiments also depend on behavioral variables, making scaling up efforts more challenging as it requires an extrapolation of behavior observed at an individual level. While incentives can be used to regulate behavior to some extent, mismatch of expectations or disagreement between stakeholders can impede the scaling up process despite successful results at the local level (Vreugdenhil 2010). If the experiments are challenging an established water management regime by suggesting innovative policy solutions and alternate pathways for resource management and transitions, collaboration between the key stakeholders is critical to break policy inertia and system lock-ins (Zhou et al. 2013). The interplay of political aspects of policy experiments with other relevant factors, are not very well researched.

The Qualitative Comparative Analysis of fifteen pilot initiatives in multiple sectors, including water resource management, aims at addressing this gap and enables us to draw some lessons for experimentation in the water sector. The presence of political support was found to be necessary for scaling up in 97 % of the cases studied, followed by the need for synergies between the experiment and ongoing policies and programmes. When found in combination with effective pilot planning and strong monitoring and evaluation, both these factors create a sufficient condition for successful scaling up in nearly 60 % of the cases studied.

References

Akumu OAK, Appida PO (2006) Privatisation of urban water service provision: the Kenyan experiment. Water Policy 8(2006):313–324

Aubin D, Varone F (2013) Getting access to water: property rights or public policy strategies? Environ Plann C: Gov Policy 31(1):154–167

Cabinet Office (2003) Trying it out: the role of 'pilots' in policy-making. Report of a Review of Government Pilots. Strategy Unit, Government of the United Kingdom

Callander S (2011) Searching for good policies. Am Polit Sci Rev 105(04):643–662, November 2011
Farrelly M, Brown R (2011) Rethinking urban water management: experimentation as a way forward? Glob Environ Chang 21:721–732
Gillespie S (2004) Scaling up community-driven development: a synthesis of experience. Int Food Policy Res Inst
Hartmann A, Linn JF (2007) Scaling Up: a path to effective development. 2020 focus brief on the World's poor and hungry people. IFPRI, Washington
Jones K (2004) Mission drift in qualitative research, or moving toward a systematic review of qualitative studies, moving back to a more systematic narrative review. Qual Rep 9(1):95–112
Kent R (2008) Using fsQCA: a brief guide and workshop for fuzzy-Set qualitative comparative analysis. University of Stirling, United Kingdom
Kirkevold M (1997) Integrative nursing research – an important strategy to further the development of nursing science and practice. J Adv Nurs 25:977–984
Kosamu IBM (2014) Conditions for Sustainability of the Elephant Marsh Fishery in Malawi. Sustainability 6: 4010–4027
Lempert RJ, Popper SW, Bankes SC (2003) Shaping the next One hundred years: New methods for quantitative, long-term policy analysis. RAND, Santa Monica
Margerum R (2012) Integrated water resources management in the United States: The rogue and willamette river cases. In Warner J, Warner JF, van Buuren A, Edelenbos J (eds) Making space for the river: governance experiences with multifunctional river flood management in the US and Europe, IWA Publishing
McFadgen B (2013) Learning from policy experiments in adaptation governance. Paper presented at the international conference on public policy. Grenoble, France
Moore M, von der Porten S, Plummer R, Brandes O, Baird J (2014) Water policy reform and innovation: a systematic review. Environ Sci Policy 38:263–271
Nastar M (2014) What drives the urban water regime? an analysis of water governance arrangements in Hyderabad, India. Ecol Soc 19(2):57. doi:10.5751/ES-06570-190257
O'Brien K, Pelling M, Patwardhan A, Hallegatte S, Maskrey A, Oki T, Oswald-Spring U, Wilbanks T, Yanda PZ (2012) Toward a sustainable and resilient future. In: Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner G-K, Allen SK, Tignor M, Midgley PM (eds) Managing the risks of extreme events and disasters to advance climate change adaptation, A special report of working groups I and II of the intergovernmental panel on climate change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp 437–486
Pattanayak SK, van den Berg C, Yang J-C, Van Houtven G (2006) The use of willingness to pay experiments : estimating demand for piped water connections in Sri Lanka. Policy, Research working paper; no. WPS 3818. Washington, DC: World Bank. http://documents.worldbank.org/curated/en/2006/01/6530300/use-willingness-pay-experiments-estimating-demand-piped-water-connections-sri-lanka
Ragin C (2006) Set relations in social research: evaluating their consistency and coverage. Political analysis advance access, Oxford University, pp 20
Ragin C (2008) Qualitative comparative analysis using fuzzy sets (fsQCA). In Rihoux B, Ragin C (eds) Configurational comparative analysis, Sage Publications, 2007
Rondinelli D (1993) Development projects as policy experiments: an adaptive approach to development administration, 2nd edn. Routledge, New York
Rotmans J, Kemp R, van Asselt MBA (2001) More evolution than revolution. Transit Manag Public Policy Foresight 3(1):15–31
Sabel C, Zeitlin J (2012) Experimentalist governance in David levi-faur. In: The oxford handbook of governance. Oxford University Press, Oxford, pp 169–83
Simmons R, Fajans P, Ghiron L (eds.) (2007) Scaling up health service delivery: from pilot innovations to policies and programmes. WHO, 2007
Spicer N, Bhattacharya D, Dimka R, Fanta F, Mangham-Jefferies L, Schellenberg J, Tamire-Woldemariam, Walt G, Wickremasinghe D (2014) Scaling-up is a craft not a science': catalysing scale-up of health innovations in Ethiopia, India and Nigeria. Soc Sci Med 121:30–38
Stoker G (2010) Translating experiments into policy. Ann Am Acad Polit Soc Sci 2010 628:47
Turton C, Bottrall A (1997) Water resource development in the drought-prone uplands. Overseas development institute. Number 18
van der Heijden J (2013) Experimentation in policy-design: insights from the building sector. Policy Sci. doi:10.1007/s11077-013-9184-z
Vreugdenhil H (2010) Pilot projects in water management: practicing change and changing practice. PhD thesis, Delft University, The Netherlands, pp 280

Vreugdenhil H, Frantzeskaki N, Taljaard S, Ker Rault P, Slinger J (2009) The next step in policy transitions: diffusion of pilot projects. 13th annual conference of the international research society for public management. Copenhagen Business School, Denmark, April 6th – 8th, 2009

Walker W, Marchau V, Swanson D (2010) Addressing deep uncertainties using adaptive policies. Technol Forecast Soc Chang, 77(6) Special Section 2

Zhou Q, Quitzau M-B, Hoffmann B, Karsten A-N (2013) Towards adaptive urban water management: Up-scaling local projects. Int J Archit Eng Constr 2(1):63–72, March 2013