Climate and Environmental Governance Network

‘Voluntary programmes for building retrofits: Opportunities, performance and challenges’

Jeroen van der Heijden

REGULATORY INSTITUTIONS NETWORK (RegNet)
The RegNet Research Paper Series is published electronically by the Regulatory Institutions Network (RegNet) at the Australian National University. Please visit our website to find out more about the work we do at regnet.anu.edu.au

The Climate and Environmental Governance Network (CEGENET) engages in theoretical, empirical and interdisciplinary research on institutions of regulation and governance (such as property rights, markets, trade regimes, treaties, national laws) and organisational actors (like states, governments, international organisations, corporations, NGOs) and the ways in which relationships and influences amongst them affect the capacity of societies to respond to environmental change and crisis.

Find out more about CEGNET at regnet.anu.edu.au/cegnet/home

© Copyright is held by the author or authors of each research paper. Electronic and paper copies may be made of a research paper, but its format may not be altered in any way without the author’s permission.

**Note:** The views expressed in each paper are those of the author or authors of the paper. They do not represent the views of the series editor, RegNet, or the Australian National University.

**Series Editor:** Dr Jeroen van der Heijden

Please forward any feedback, submissions or requests to: Regnet@anu.edu.au
‘Voluntary programmes for building retrofits: Opportunities, performance and challenges’

Jeroen van der Heijden, The Australian National University; University of Amsterdam.

Working Paper

Citation

This paper can be cited as: Van der Heijden, J. 2014. Voluntary programmes for building retrofits: Opportunities, performance and challenges. CEGNet Working Paper, No. 48, Regulatory Institutions Network.

Abstract

Around the globe governments, businesses and citizens are actively involved in voluntary programmes that seek an improved uptake of retrofits of the existing building stock. Using fuzzy set qualitative comparative analysis (fsQCA) this article seeks to understand the opportunities, performance and constraints of such programmes. Building on a series of 20 voluntary programmes in Australia, the Netherlands and the United States (including a series of 101 original interviews) it finds that the majority of these have not succeeded in incentivising their participants to take meaningful action. The article provides insight into why the majority of these programmes have underperformed, and what binds together the small number of programmes that have achieved positive results.

Keywords

Urban transformation and adaptation, urban sustainability and resilience, retrofitting, collaborative governance, voluntary programmes
1 Introduction
Chapter 8 of the recent IPCC Fifth Assessment Report, *Urban Areas*, stresses the importance of transforming and adapting urban areas to a changing climate (IPCC, 2014). The report acknowledges that key to a successful transformation of urban areas is the retrofitting of existing buildings. After all, at any time existing buildings make up the vast majority of our built environment. The report is, however, limited in its suggestions of how to govern the transformation of existing buildings. Among the few suggestions is a call for increased regulatory requirements for existing buildings, which *de facto* implies a call for mandatory retrofits. Governments around the globe will, however, not find it easy to introduce increased requirements for existing buildings, or to mandate retrofits.

Under a current ‘business as usual’ regulatory strategy existing buildings and infrastructures are normally exempted from new and amended construction regulation. This is known in the (legal) literature as ‘grandfathering’ (Nash & Revesz, 2007; Vinagre Diaz, Wilby, & Belén Rodríguez González, 2013). In effect, grandfathering reduces the transformative capacity of regulation and legislation. Urban areas (in developed economies) normally grow and transform at less than 2% per year. It may therefore take 40 to 70 years for new regulations to transform all buildings and infrastructure in cities (IEA, 2009).

At the same time questions have arisen as to the effectiveness of ramping up regulatory requirements for existing buildings and mandatory retrofits. The existing building stock in most Western societies may be too varied in age, quality, and condition (Hassler, 2009; Hassler & Kohler, 2014) to allow for a ‘one size fits all’ solution to such increased regulatory requirements (Galvin, 2014; Sunikka-Blank & Galvin, 2012). Such increased regulatory requirements and mandatory retrofits further do not respond to other problems that stand in the way of the transformation of existing buildings. This includes, but is not limited to, split incentives between building owners and building users (Gillingham, Harding, & Rapson, 2012), an unwillingness of banks to provide mortgages for retrofits (Managan, Layke, Monica, & Nesler, 2012), or sub-optimal behaviour by building occupants (Osbaldeston & Schott, 2012).

Policy makers, businesses and NGOs increasingly acknowledge the need for a focus on existing buildings in the transformation of urban areas. They also acknowledge the difficulties of introducing increased regulatory requirements or mandatory retrofits for existing buildings (EC, 2013; IEA, 2012). Seeking to overcome these problems governments around the globe are trialling innovative non-regulatory and often non-mandatory interventions (Van der Heijden, 2014). More and more they do so by engaging directly with businesses and citizens in developing voluntary programmes that encourage the voluntary retrofitting and upgrading of existing buildings (Gollagher & Hartz-Karp, 2013; Hoffmann, 2011; Potoski & Prakash, 2009).

It remains, however, in question whether voluntary programmes are successful in achieving a meaningful transformation of existing buildings. It also remains in question what design characteristics of these programmes may add to their successful outcomes. This then is the aim of the current article. It studies a sample of 20 voluntary programmes that aim to improve the environmental and resource sustainability of existing buildings, or their resilience to
hazards in Australia, the Netherlands, and the United States. The article applies fuzzy set comparative qualitative analysis (fsQCA) to better understand what design conditions of these programmes are related to their (positive or negative) outcomes.

2 Voluntary programmes: a brief review of the literature
Over the last decades voluntary programmes have gained increasing scholarly and policy interest. Individuals and organisations participating in these programmes pledge to change their behaviour in such a way as to create desired societal outcomes beyond what is required by governmental regulation. They can best be understood as rule regimes that give exclusive rewards to their participants, such as the branding of their goods and services, or showcasing industry leadership (Potoski & Prakash, 2009). Voluntary programmes have gained particular prominence in the addressing of complex environmental risks, including climate change (Borck & Coglianese, 2009; Potoski & Prakash, 2013).

Well-known examples from the built environment are the international building assessment programmes LEED (Leadership in Energy and Environmental Design) and BREEAM (BRE Environmental Assessment Methodology). These voluntary programmes seek to improve the environmental sustainability of buildings by making visible their environmental performance. To do so, buildings are assessed against a set of sustainability requirements and their performance is certified in a particular class. In this way buildings can be compared according to their relative score – i.e., for a building developer, property owner, or tenant it is easy to understand that a ‘platinum’ rated building has a better environmental performance than a ‘bronze’ rated one (Yudelson & Meyer, 2013). Programmes such as LEED and BREEAM are widely implemented around the globe and normally seek to push the performance of buildings beyond national construction codes (Cole & Valdebenito, 2013). For building developers, building owners and building occupants, there are clear incentives to have their buildings certified under such voluntary programmes. In doing so, they can make visible their environmental credentials, which adds to their overall social corporate responsibility image (Dixon, Ennis-Reynolds, Roberts, & Sims, 2009). Early research also indicates that these certified buildings yield higher rents and higher sales prices than non-certified buildings (Eichholtz, Kok, & Quigley, 2010).

The literature on voluntary programmes highlights that these are applied in a wide range of contexts (countries and sectors); and that voluntary programmes come in many forms, building on different designs (OECD, 2003). The literature further indicates that not all voluntary programmes are equally successful in achieving their goals, i.e., creating desired societal outcomes (for extensive reviews of the literature, see Borck & Coglianese, 2009; Prakash & Potoski, 2012; Van der Heijden, 2012). It considers that both contextual conditions and design conditions of voluntary programmes influence their outcomes. In terms of contextual conditions, it considers the following as most relevant in influencing the performance of voluntary programmes:

- Existing legislation: it goes without saying that existing (environmental) legislation should allow for voluntary programmes to be developed (Kollmuss & Agyman, 2002), and provide enough niches for voluntary programmes to fill (NeJaime, 2009). More importantly, existing legislation may affect the performance of these programmes. Different arguments hold. For instance, if existing environmental legislation sets relatively high standards in a particular country, its citizens (individuals and
organisations alike) may be more aware of environmental risks and the need to act on these than under a situation of relatively low standards of environmental legislation. This may make them more willing to participate in voluntary programmes (Bressers, De Bruijn, & Lulofs, 2009). Also, under the threat of increased mandatory legislation individuals and organisations may seek to join voluntary programmes hoping that this will forestall the implementation of future legislation (Short & Toffel, 2010).

- **Economic circumstances:** it is often assumed that higher levels of wealth coincide with higher levels of environmental concerns (Givens & Jorgenson, 2013). In other words, the higher the disposable income of consumers (individuals and organisations alike) the more likely they will demand products or services that in one way or another have a less negative impact on the natural environment (Baron & Diermeier, 2007). Through participation in voluntary programmes producers may seek to tap into this market of sustainable consumers.

- **Societal pressure:** related to the above contextual conditions, it is often argued that under a situation of high societal pressure for sustainable practice individuals and organisations are more likely to participate in voluntary programmes. These programmes may provide them with a way of seeking public recognition for their beyond regulatory compliance performance (Briscoe & Safford, 2008).

In terms of design conditions, the current literature considers the following as most relevant in influencing the performance of voluntary programmes:

- **Rewards for participants of a voluntary programme:** a major motivation for (prospective) participants to join a voluntary programme is the financial gain they may achieve. For instance, participants may be able to lease their retrofitted buildings at higher rates, or they may see reduced energy costs as a result of better insulated buildings. It is expected that the higher the gain for participants, the higher the outcomes a programme achieves (Croci, 2005). But also non-monetary gains are considered of relevance. For instance, in joining a voluntary programme participants may obtain information on how to improve the energy performance of their buildings, or may build close networks with policymakers and peers. A final reward that the literature considers of relevance is the ability to showcase leadership through participation in a voluntary programme. Participants may seek exposure for their leadership and be recognised as such. It is expected that the more leadership is recognised and rewarded, the higher the outcomes a programme achieves (Borck & Coglianelse, 2009).

- **Stringency of a voluntary programme:** the current literature considers that both the internal rules of a programme as well as their enforcement influence the performance of voluntary programmes. A voluntary programme may require participants to move their performance well beyond the requirements set by governmental building regulation, or it may ask for marginal beyond compliance behaviour. It is expected that the stricter the participation criteria of a voluntary programme, the lower the outcomes a programme achieves because participation in the programme asks for considerable effort from (prospective) participants (Potoski & Prakash, 2009). In terms of enforcement, a voluntary programme may require self-monitoring by participants or third party certification as proof of compliance. It is expected that the stricter this enforcement of participation criteria, the lower the outcomes because more non-compliant behaviour is identified (Potoski & Prakash, 2009).
• **The role of government in a voluntary programme**: last but not least, the current literature on voluntary programmes suggests increasing awareness of the role of governments in voluntary programmes. The more governments are involved in the development and administration of a voluntary programme the more credibility it may have in the eyes of (prospective) participants, and the lower their (financial) risks when joining the programme. It is expected that the more government involvement in a voluntary programme, the higher the outcomes a programme achieves (Gunningham, 2009). Also, governments may play a major role as ‘client’ of a programme. Governments are a major ‘consumer’ of buildings. As participants or launching customers of voluntary programmes they may influence these programmes’ performance. It is expected that the more government participation in a voluntary programme, the higher the outcomes a programme achieves (Hofman & De Bruijn, 2010).

Whilst these conditions are often identified as being related to the performance of voluntary programmes, it remains unclear how they relate exactly. Moreover, earlier empirical studies have found that voluntary programmes with similar designs such as pay-per-plastic-bag fees (Ackerman, 1997), organic food labelling (Thøgersen, 2010), building assessment classification and certification (Fowler & Rauch, 2006), and revolving loan funds (Boyd, 2013) show different outcomes depending on how their design conditions interact with contextual conditions (e.g., existing legislation, economic circumstances; Borck & Coglianese, 2009). Furthermore, some studies indicate that a single design (e.g., building assessment classification and certification) implemented in a number of similar contexts (e.g., the United States, Australia, the United Kingdom) may nevertheless result in different outcomes due to the role of governmental actors in these arrangements (Fowler & Rauch, 2006). This all indicates that the outcomes of voluntary programmes are likely caused by different interacting conditions (i.e., conjunctural causation), and that different (configurations of interacting) conditions may cause a similar outcome (i.e., equifinality).

3 **Research design: case selection, data collection, and data analysis**

To gain insight into the opportunities, performance and challenges of voluntary programmes that seek to achieve retrofits of existing buildings, a series of 20 programmes (‘cases’) from Australia, the Netherlands and the United States were studied. The focus of the study is to gain insight into the design conditions of the voluntary programmes under scrutiny. The countries are chosen because they show considerable similarity on the contextual conditions that the current voluntary programme literature considers to be related to the outcomes of such programmes. They show considerable similarity in their statutory building code regimes (Liu, Meyer, & Hogan, 2010), and particularly their focus on resource sustainability and energy consumption (IEA, 2013). The countries rank fairly comparably in terms of economic development and citizens’ standards of living (UNDP, 2013). The countries further rank fairly comparably in terms of environmental awareness of citizens and businesses (OECD, 2013). Finally, in all countries the retrofitting of existing buildings may be considered the ‘low hanging fruit’ in terms of reduced carbon emissions and energy savings (IPCC, 2014), whilst in all countries studied the existing building stock will be considerably affected by a changing
climate. This all motivates the development of voluntary programmes in these countries, and may be expected to motivate individuals and organisations to join these.

3.1 Case selection
Cases were selected from a larger pool of cases derived from an extensive internet search using search terms such as ‘building AND retrofit’ and ‘existing building AND sustainability’. From this pool (further discussed in Van der Heijden, 2014) cases were selected for further study (8 cases for Australia, 6 for the Netherlands, and 6 for the United States) when they met a number of criteria (i.e., a stratified sample):

- They explicitly focus on increasing the environmental and resource sustainability of existing buildings through retrofits of existing buildings.
- They all set requirements that ask property owners to voluntarily make changes to their buildings well beyond requirements laid down in building legislation and regulation.
- Cases were selected to include a variety of approaches to goal achievement. It was expected that including different approaches to goal achievement (i.e., different designs of the programmes) helps to better understand what design conditions matter to achieve positive outcomes for voluntary programmes. These are:
  - Collaborative networks that aim to learn how building retrofits can be achieved without the use of statutory regulation (4 cases).
  - Building assessment programmes (4 cases).
  - Innovative forms of financing that help owners of existing buildings acquire funding for retrofits (7 cases); this is necessary since many banks are often not willing to fund such retrofits (e.g., Pivo, 2010).
  - Voluntary programmes that target a particular regulatory barrier (e.g., regulation that hampers the instalment of solar panels on existing strata buildings; 5 cases).
- Only cases were selected that have matured to at least two years of actual implementation – i.e., it was expected that some time is needed for the cases to achieve outcomes.

Whilst space limitations do not allow for an extensive discussion of each of the programmes studied, Table 1 gives a brief overview of these.

TABLE 1 ABOUT HERE

3.2 Data collection
In order to understand the development, implementation and performance of the cases under analysis these were studied intensively. Existing data on the cases was collected from programme websites, programme reports, and other documentation. Novel data on the cases were obtained through a series of in-depth face-to-face interviews carried out in 2012 and 2013. Interviewees were traced through internet searches and through social-network websites, particularly LinkedIn. In addition, prospective interviewees were identified through a form of targeted snowball sampling, where interviewees were asked to provide names of other key-stakeholders in a programme. Interviewees were targeted for their in-depth knowledge with one or more cases under scrutiny. This resulted in a pool of 101 interviewees (55

---

1 Please note, none of the country contexts clusters together in any of the solution formulae of the fsQCA analysis (see section 4). This indicates that the country contexts are appropriately constant for the current study.
Australian\(^2\), 27 from the Netherlands, 18 from the United States) from various backgrounds; Table 2 gives an overview.

TABLE 2 ABOUT HERE

The interviews were based on a semi-structured questionnaire that provided a structure of checks and balances to assess the validity of the findings. Also, the interviews were recorded and transcribed into a report that was sent back to the interviewees for validation. The interviewees were often aware of and involved in more than one case. It is expected that this (partly) helped to overcome a sampling bias of administrators (and participants) who were overly enthusiastic about their ‘own’ case (Sanderson, 2002).

3.3 Data analysis method: fsQCA

The data were processed by means of a systematic coding scheme and qualitative data analysis software (Atlas.ti). By using this approach the data were systematically explored and insights were gained into the ‘repetitiveness’ and ‘rarity’ of experiences shared by the interviewees, and those reported in the existing information studied. This allowed in-depth understanding of the individual cases, and it may further assist in tracing across-case patterns (Venesson, 2008). The data were further analysed using fuzzy set qualitative comparative analysis (fsQCA) logic, techniques and FS/QCA software (version 2.5). Following on from the extant literature on voluntary programmes (see above) QCA was chosen as a data analysis methodology because it allows for ‘unravelling causally complex patterns in terms of equifinality [and] conjunctural causation’ (Schneider & Wagemann, 2012, 8).

QCA differs from other data analysis methods in its focus. QCA is grounded in set theory, a branch of mathematical logic that allows studying in detail how causal conditions contribute to a particular outcome. ‘The key issue [for QCA] is not which variable is the strongest (i.e., has the biggest net effect) but how different conditions combine and whether there is only one combination or several different combinations of conditions (causal recipes) of generating the same outcome’ (Ragin, 2008, 114). QCA helps to trace patterns of association between these conditions in a highly systemised manner and allows for systematic comparisons between empirical observations (i.e., cross-case), whilst allowing for in-depth within-case understanding of the individual observations (Rihoux & Ragin, 2009).

QCA may be understood as an approach that helps researchers to come to evidence based typologies from their data (cf., Fiss, 2011). For instance, this article presents a typology of program designs that are related to successful outcomes, and a typology of programs that are related to unsuccessful outcomes. Within QCA terminology the individual types are referred to as ‘paths’, and a full typology is referred to as ‘solution’. It should be kept in mind that whilst QCA uses numerical symbols, it is a qualitative method. The numerical information provided throughout this article should be understood as descriptions of data patterns that underlie the dataset, but not as a simplistic reduction of the qualitative data obtained (cf., Schneider and Wagemann, 2012). For instance, a ‘consistency score’ helps the researcher to understand how well a path and the full solution reflects the empirical data, whilst a ‘coverage score’ helps

\(^2\) The study was started in Australia. Interview accounts from the Australian cases largely confirmed the extant theory on voluntary programmes, and pointed out that for the purpose of the current study much data is available on programme websites and in programme reports. The focus of the interviews in the Netherlands and the United States was to validate data from these sources and to fill in gaps in the data. This explains why a relatively larger number of participants per case were interviewed in Australia than in the two other countries.
QCA has since the mid-1990s quickly evolved as an accepted research practice for the type of study presented in this article, and has been applied in hundreds of studies in the humanities (Ragin, 2008; Rihoux, 2013; Rihoux & Ragin, 2009; Schneider & Wagemann, 2012). It should be noted, however, that this is one of the first studies that applies QCA to better understand the performance of voluntary programmes.

3.4 Outcomes and conditions of interest

The aim of the research presented is to understand how particular design conditions of voluntary programmes relate to their outcomes. The extant literature on voluntary programmes considers at least two outcomes relevant in the assessment of their performance (Borck & Coglianese, 2009; Potoski & Prakash, 2009):

- O1: The number of participants a voluntary programme attracts. For the current study it was considered how well a programme has performed in achieving its stated ambitions in terms of attracting participants.
- O2: The contribution of a programme to a desired societal end. For the current study it was considered how well a programme has performed in achieving its stated ambitions in terms of building retrofits.

Building on the extant literature on voluntary programmes (see Section 2), the current study considers the following design conditions of interest to better understand the performance of the 20 voluntary programmes under scrutiny:

- Rewards for participants of a voluntary programme: financial gain for participants (Fg); non-monetary gain for participants (Nm); and, showcasing leadership by participants (Le).
- Stringency of a voluntary programme: participation criteria (Pc); and, enforcement of these criteria (En).
- The role of government in a voluntary programme: government involvement in the development and administration of a voluntary programme (Gi); and, government involvement through participation as participant or launching customer of a voluntary programme (Gp).

4 Data analysis

Table 3 presents the data for this study. The data observations are scored on a four point scale to indicate the comparative (qualitative) differences in observations.

TABLE 3 ABOUT HERE

3 Understanding that fsQCA may be a less well-known method to some of the readers an online appendix has been prepared for the interested reader (building on: Ragin, 2008; Schneider & Wagemann, 2012). The supplementary online appendix pays in-depth attention to the logic underlying fsQCA. This appendix further gives a step-by-step description of how fsQCA has been applied in this study.

4 Please note the supplementary file, step 4, gives extensive insight into the operationalisation and calibration of the conditions and outcomes (the data observations).
This table presents an important insight: whilst a majority of the programmes (70%; n=15) have attracted a substantial number of participants, or even the expected number of participants, substantially less (30%; n=5) have performed as well in achieving retrofitted buildings. This supports earlier empirical studies on voluntary programmes. Potoski and Prakash (2009), among others, have earlier found that high numbers of participants in a programme is no guarantee that it also performs well in achieving other outcomes. The remainder of this article will therefore focus on what binds together the handful of programmes that do show hopeful results in terms of making their participants act towards the goal of the programme (i.e., achieving building retrofits) and what binds together the large number of programmes that have not achieved such results.

4.1 Necessary conditions
Following fsQCA practice the data are first analysed for necessary conditions before being exposed to more complex analysis to identify (configurations of) sufficient conditions (Rihoux & Ragin, 2009, Chapter 5, Box 8.1; Schneider & Wagemann, 2012, Chapter 11). Table 4 presents the results of this analysis for necessity.\(^5\)

TABLE 4 ABOUT HERE

Conditions should only be considered necessary if their consistency scores\(^6\) are very high; a cut-off point of 0.90 is advised (Rihoux & Ragin, 2009, 45). As can be seen from Table 3, only the condition *financial gain for participants* has a consistency score of 0.95. However, the low coverage score\(^7\) of 0.35 indicated that this is likely a trivial necessary condition in achieving this outcome (Schneider & Wagemann, 2012, 232-237). All other conditions do not meet the consistency score of 0.90.

In sum, the data do not point to any distinct (relevant) necessary condition for causing the outcome O2 (building retrofits). Because the existing literature considers such a variety of conditions that may be related to this outcome, and because none of the distinct conditions appears a (relevant) necessary condition, it is likely that: (i) different conditions cause a similar outcome (i.e., equifinality); and that (ii) conditions interact in causing the outcomes (i.e., conjunctural causation). This is what the following sections seek to understand.

4.2 Sufficient conditions for achieving substantial to high numbers of building retrofits
This section seeks to better understand what binds together the handful of programmes that have resulted in achieving a substantial to high number of retrofitted buildings (as compared to stated ambitions). The data are analysed to logically reduce the empirically observed configurations (Rihoux & Ragin, 2009, Chapter 5, Box 8.1; Schneider & Wagemann, 2012, Chapter 11). That is, data are studied to gain insight into (configurations of) conditions of voluntary programmes that may be sufficient to cause substantial to high numbers buildings retrofitted. Table 5 gives a summary of the findings.\(^8\)

TABLE 5 ABOUT HERE

---

\(^5\) For a discussion of the analysis, see the supplementary file, step 6.

\(^6\) Consistency indicates how strongly the condition relates to the outcome.

\(^7\) Coverage indicates how relevant the condition is for causing the outcome.

\(^8\) For a discussion of this analysis, see the supplementary file, steps 7 to 9.
Table 5 adopts a notation and presentation of causal configurations ('solutions') that are sufficient for causing the outcome of interest as introduced by Ragin (2008, 205) and applied by others (Erkens & Van der Stede, 2013; Fiss, 2011). Large full circles (●) indicate core causal conditions that must be present in causing the outcome; large crossed out circles (ø) indicate core causal conditions that must be absent. Small full circles (●) indicate contributing causal conditions that must be present to cause the outcome; small crossed out circles (ø) indicate contributing causal conditions that must be absent.9

The analysis first considered the distribution of cases (empirical observations) across all logically possible configurations of conditions (i.e., a maximum of $2^7=128$ combinations is possible for the seven conditions derived from the extant literature). Following Ragin (2008) only configurations with at least one observation were considered because of the relatively small number of cases studied – which de facto implies that all the empirical data collected were studied. Of the possible 128 configurations of conditions 14 were observed (with some configurations being observed in more than one case, e.g., cases 14, 24 and 42 are characterised by a similar configuration of conditions, see also Table 2). Following Ragin (2008, 142-144) a consistency score of $>0.75$ was chosen to distinguish configurations that are subsets of the outcome from those that are not. Four configurations met this requirement.

Coverage scores in the tables are a measure to indicate the importance of (configurations of) sufficient conditions (Schneider & Wagemann, 2012, section 5.3). For instance, a high coverage score indicates that a configuration is of high empirical importance in reaching the outcome under scrutiny (here a substantial to high number of building retrofits). The overall coverage solution for this analysis (0.65) may be considered substantial (Ragin, 2008). Because of the possibility that individual cases can be characterised by more than one (simplified) solution it is also of interest to know how much of the outcome is covered by a solution, and how much of the outcome is exclusively covered by a specific solution. This is what the scores for, respectively, Raw coverage and Unique coverage indicate. Here solution ‘2’ covers more of the outcome than solution ‘1’ and solution ‘3’.

Consistency scores give insight into the degree to which the configuration relates to the outcome: the higher the consistency score, the stronger the evidence that the configuration relates to the outcome under scrutiny. A score of $>0.75$ is advised as a cut-off point when interpreting solutions (Ragin, 2008, 144; Schneider & Wagemann, 2012, section 5.2). Whilst the Overall solution consistency of this analysis (0.79) indeed passes that threshold, the consistency scores of the individual solutions may raise some concerns about how strongly particularly solution ‘3’ is related to the outcome and is therefore not further analysed in this article.

The two solutions that meet the threshold can be read as:

- **Solution 1**: Voluntary programmes that are characterised by a high financial gain for building retrofits; absent of strong participation criteria and enforcement of these, and absent of government involvement in the programme.

- **Solution 2**: Voluntary programmes that are characterised by strong government involvement in the programme, combined with strict enforcement of otherwise low participation criteria, and that provide high financial gain for building retrofits.

---

9 Fiss (2011) provides a highly accessible discussion of this notation.
These solutions are further interpreted in section 5, after addressing the voluntary programmes that have been less successful in achieving high numbers of retrofitted buildings.

4.3 Sufficient conditions for an absence or only marginal production of building retrofits

This section seeks to better understand what binds together the majority of arrangements that have performed less well, and have not achieved any or only a marginal number of building retrofits (as compared to stated ambitions). The same approach is followed as in the previous section, except here the negated outcome O2 (building retrofits) is studied (cf., Ragin, 2008) – i.e., a focus on the “-” and “-” scores for this outcome in Table 2. Table 6 gives a summary of the findings.10

TABLE 6 ABOUT HERE

Table 6 indicates that five solutions are related to the outcome of interest (here: an absence or only marginal number of building retrofits). The solution consistency (1.00) and solution coverage (0.82) may both be considered high – i.e., the overall solution strongly relates to the outcome, and the solution is of high empirical importance in reaching the outcome (cf., Schneider & Wagemann, 2012). The solution consistency and solution coverage of all solutions is sufficient.

The five solutions can be read as:

- **Solution 4**: Voluntary programmes that are characterised by strict participation criteria; absent of high financial gain for building retrofits, and absent of government involvement in the programme.
- **Solution 5**: Voluntary programmes that are characterised by strict participation criteria that are strictly enforced, absent of government involvement in the programme.
- **Solution 6**: Voluntary programmes that lack the rewarding of leadership, do not provide non-monetary gains, are absent of government involvement in the programme, but that face strict enforcement of their participation criteria.
- **Solution 7**: Voluntary programmes that are characterised by a high financial gain for building retrofits, but that lack other incentives for participation (such as a rewarding of leadership and non-monetary gains) and face strict participation criteria.
- **Solution 8**: Voluntary programmes that are characterised by strong government involvement in the programme, but that lack a rewarding of leadership and lack government participation, combined with lenient enforcement of these programmes’ participation criteria.

5 Discussion: opportunities and constraints

What is of particular interest about the set of 20 programmes studied in this article is the wide variety of programmes (see Table 1): they address different types of buildings (e.g., commercial and residential), different types of building owners (e.g., commercial property owners, housing associations, and home owners), and seek to overcome problems that stand in the way for building retrofits in different ways (e.g., building assessment programmes, innovative forms of financing, and collaborative networks). They seek an ongoing adaptive change of the existing building stock.

---

10 For a discussion of this analysis, see the supplementary file, steps 7 to 9 and step 10.
This all seems to indicate that these 20 voluntary programmes are highly tailored to their local context and the individuals and organisations in this context. Such tailoring may be an answer to the oft expressed critique to increased regulatory requirements for existing buildings and mandatory retrofits (Galvin, 2014; Hassler, 2009; Hassler & Kohler, 2014; Sunikka-Blank & Galvin, 2012), and may be an answer to other problems that stand in the way of building retrofits, such as split incentives or an unwillingness of banks to provide mortgages for retrofits (Gillingham et al., 2012; Managan et al., 2012). Voluntary programmes may be a hopeful alternative to a slow moving ‘business as usual’ scenario of regulatory requirements that seek to transform and adapt urban areas to a changing climate. The current study does, however, also point out a number of constraints of voluntary programmes.

5.1 The difficulty of incentivising participants to retrofit buildings
A main finding of the current study is that a high number of participants in a voluntary programme for building retrofits are by no means a guarantee for achieving a large number of retrofits. The seven types of programmes identified in sections 4.2 and 4.3 are of interest. The two solutions related to substantial to high numbers of building retrofits (as compared to stated ambitions) produced by the participants in the voluntary programmes (section 4.2) indicate that participants are likely to improve the performance of their buildings if doing so is not too complicated, and results in monetary or non-monetary gains. That is, in ‘solution 1’ participants may seek cost-savings through reduced energy consumption or tapping into a new consumer market for highly sustainable buildings. In ‘solution 2’ participants may seek to gain from the strong government involvement in these programmes. Such government involvement may make these programmes more credible in the eyes of the wider public. Government involvement may further take away the financial or administrative risks for participants.

The five solutions related to not achieving any or only a marginal number of building retrofits (as compared to stated ambitions) produced by the participants (section 4.3) indicate that participants are unlikely to improve the performance of their buildings if the cost of effort of doing so does not outweigh the gains. Four solutions (solutions 4 to 7) are characterised by overall low rewards for participants combined with substantial effort to achieve such rewards, see in particular the clustering of core causal conditions in the stringency of the arrangements in these solutions (see Table 5).

These insights (related to solutions 1 and 2, and 4 to 7) were shared by the wide range of interviewees; Table 7 presents some direct quotes from the interviews.

TABLE 7 ABOUT HERE

The last solution, solution 8, deviates somewhat from the other solutions that are related to not achieving any or only a marginal number of goods and services (solutions 4 to 7). This solution seems to indicate that voluntary programmes with this specific configuration of design conditions lack focus: whilst governmental actors are involved, the programmes do not have a clear focus on specific rewards. Voluntary programmes following this design may need more in-depth exploration in future research.

5.2 Limited success for voluntary programmes for building retrofits
Combined, these six solutions (solution 1 and 2, and 4 to 7) paint an insightful picture of the ability of voluntary programmes to overcome problems that stand in the way of building retrofits in the transformation of urban areas. That is, the current research indicates that it should be made very clear to (prospective) participants of voluntary programmes that the benefits of these programmes (i.e., financial gain, administrative support, having their leadership recognised), outweigh the costs of participating and acting towards the goals of these programmes. Such programmes may then be expected to achieve positive outcomes in terms of building retrofits. However, at the same time not too much should be asked from participants. The current research also indicates that programmes that set high participation criteria are unlikely to achieve successful outcomes.

Exactly because of this combined set of insights not too much should be expected from voluntary programmes for building retrofits. If no high participation criteria can be set, then programmes are unlikely to stimulate participants to take far reaching action (cf., Borck & Coglianese, 2009; Potoski & Prakash, 2009). Furthermore, if only high gains make participants act towards the goal of these programmes, then it is unlikely that large groups of building owners will wish to be involved in such voluntary programmes. For instance, showcasing leadership can be highly attractive to large corporations since it adds to their social corporate responsibility strategies (Dixon et al., 2009). Showcasing leadership is, however, unlikely to be an attractive incentive to households, or small enterprises.

Furthermore, the financial gains of energy retrofits of large office buildings may add up to tens of thousands of dollars annually for a large property owner, but at best to a few hundred dollars for a household (EIA, 2013). Such a small saving will likely be considered futile within the household’s finances (cf., Cialdini, 2009), which may take away the attractiveness of the financial rewards that many voluntary programmes build on.

These insights were also shared by a wide range of interviewees; Table 8 presents some direct quotes from the interviews.

TABLE 8 ABOUT HERE

5.3 Further opportunities, constraints, and future research
It goes without saying that voluntary programmes for building retrofits come with more opportunities and constraints than those identified in the current study. As with all empirical research, a number of caveats apply to the current study. The study has only explored voluntary programmes in Australia, the Netherlands and the United States. Therefore, the results of this study cannot be exported to other contexts (i.e., other countries or sectors) before carefully analysing what differences in the contexts may further affect the outcomes of voluntary programmes. Section 2 has pointed out a number of contextual conditions that researchers may wish to include in their future studies of voluntary programmes.

More importantly, the current study has not looked at the actual energy performance or reduced resource consumption of the building retrofits that resulted from the programmes studied. Nor has the current study considered whether the embedded energy of a building retrofit outweighs the energy savings resulting from the retrofit. It may, for instance, very well be that the retrofits achieved through the voluntary programmes studied do not achieve high levels of resource or environmental sustainability. Earlier research has, after all, indicated rebound effects where occupants of an energy efficient building do use relatively more energy resources.
than those of an energy inefficient building (Sunikka-Blank & Galvin, 2012). Only a few of the programmes studied do (publicly) provide the energy savings achieved. To gain a fuller insight into the opportunities and results of such programmes it would be advisable that all voluntary programmes in this area follow these examples; and it would be advisable that future research on this topic does include this focus.

Also, the current study has not considered other positive outcomes of voluntary programmes than those that have focal attention in the current voluntary programme literature (i.e., the study largely builds on a deductive strategy). For instance, voluntary programmes for building retrofits may result in improved functionality, aesthetics, indoor environmental quality, security, or prepare the existing building stock for (expected) increased temperatures. These may all be considered additional rewards for participants of a voluntary programme that scholars may wish to include in future studies.

Finally, the current study has indicated the value of fsQCA methodology for the type of study presented: it helps to better understand how complex and interacting conditions affect the performance of governance instruments such as voluntary programmes. The methodology could, for instance, also be applied to gain a better understanding of how contextual conditions affect the performance of a specific type of voluntary programme as applied in different countries. A logical topic would be an fsQCA study of highly popular building assessment and certification programmes such as LEED and BREEAM. Another application could be a comparative study of the performance of mandatory regulation and voluntary programmes within a similar context.

6 Conclusion
This article aimed to gain insight into whether and how voluntary programmes are effective in achieving building retrofits. It revealed that design conditions of voluntary programmes interact in causing their effects (i.e., conjunctural causation). This is a relevant insight because it indicates that the designers of such programmes should choose the mix of design conditions carefully. The potentially positive impact of one condition (e.g., financial gain for participants) may be cancelled out when combined with another particular condition (e.g., high participation criteria). The study also revealed that different configurations of design conditions may similarly affect the outcomes of voluntary programmes (i.e., equifinality). This is again a relevant insight because it indicates that the designers of such programmes can choose from a variety of designs to suit their needs and those of expected participants.

Most importantly, the study revealed that not too much should be expected from voluntary programmes in achieving large numbers of building retrofits that add to a timely transformation and adaptation of urban areas. Whilst voluntary programmes may be successful in particular niche markets (e.g., the high end of the commercial sector) they do not seem to achieve sweeping results across the board. Future research may therefore be interested to better understand whether, how, and where, voluntary programmes are a promising addition to or alternative for more coercive approaches. After all, building retrofits should become mainstream if we wish to utilize the transformative potential that the built environment holds. Voluntary programmes may be a means to that end, and should not be an end in themselves.
### Tables

**Table 1 – Overview of voluntary programmes studied**

| Name                                | Brief description**                                                                 |
|-------------------------------------|-------------------------------------------------------------------------------------|
| 1200 Buildings                      | Melbourne based tripartite financing tool that funds retrofits of existing commercial property. |
| Amsterdam Investment Fund*          | Revolving loan fund of the City of Amsterdam that issues loans to, among others, building developments and retrofits that seek to achieve high levels of environmental performance. |
| Better Building Partnership         | Partnership between the City of Sydney and local commercial property owners committed to reduce their energy consumption. |
| Better Buildings Challenge          | United States’ Government technical and administrative support for commercial and industrial building owners aiming to improve energy efficiency of their buildings; it further acts as a platform that matches financers with building owners seeking funds for retrofits. The programme is implemented with regional and local variance. |
| Billion Dollar Green Challenge      | U.S. wide programme that encourages colleges, universities, and other non-profit institutions to invest a combined total of one billion dollars in self-managed revolving funds that finance energy efficiency improvements. |
| BREEAM-NL* (BRE Environmental Assessment Method applied in the Netherlands) | Building assessment programme. Aims to stimulate developers and property owners to build and retrofit buildings with high levels of environmental performance. Originally started in the United Kingdom, now applied throughout the world including the Netherlands. |
| Building Resilience Rating Tool*    | Programme developed by the Australian insurance industry that rates the resilience of homes to common extreme weather events, and seeks to stimulate retrofits. |
| Dutch Energy Service Company        | Energy Service Companies (ESCOs) aim to reduce the energy consumption of their clients. The Dutch Government supported ESCOs by reducing the legal barriers for them to operate. |
| Energy Leap                         | Civil society to government collaboration that seeks to improve energy efficiency of the built environment. It develops pilot projects and governance experiments, financially supports leaders in the sector, and seeks to overcome regulatory barriers. |
| Environmental Upgrade Agreements    | Sydney based tripartite financing tool that funds retrofits of existing commercial property. |
| Green Deals*                        | Covenants between the Government of the Netherlands and local businesses and households committed to reduce their greenhouse gas emissions. |
| Green Finance SF                    | San Francisco adaptation of PACE programme (see below). |
| Green Star*                         | Building assessment programme. Aims to stimulate developers and property owners to build and retrofit buildings with high levels of environmental performance. |
**Table 1 – Continued**

| Programme                                      | Description                                                                                      |
|------------------------------------------------|-------------------------------------------------------------------------------------------------|
| **GreenStrata**                                | Australian wide programme that provides information to owners of units in multi-residential buildings, and their managers, on how to improve the environmental sustainability of these buildings. |
| **LEED* (Leadership in Energy and Environmental Design applied in the United States)** | Building assessment programme. Aims to stimulate developers and property owners to build and retrofit buildings with high levels of environmental performance. |
| **PACE (Property Assessed Clean Energy)**       | Tripartite financing programme that allows local governments in the United States to issue bonds to investors and use these funds as loans for energy retrofits to home-owners and commercial property owners. The programme is implemented with regional and local variance. |
| **Retrofit Chicago**                            | Chicago adaptation of the Better Buildings Challenge.                                            |
| **Small Business Improvement Fund**             | Chicago (United States) based programme that financially supports building owners to improve the energy efficiency of their buildings; funds are attracted through Tax Increment Financing (TIF) revenues. |
| **Sunny Rentals**                               | Collaborative project by housing corporations, their advocacy group, and governments in the Netherlands. It seeks to overcome legal barriers that stand in the way of the instalment of solar panels on rentals – both individual homes and condominiums. |
| **Sustainable Backyard Program**                | Collaboration between the City of Chicago, an NGO and garden material suppliers. It aims to improve the environmental sustainability of homes and gardens. |

* Please note: A number of the voluntary programmes studied focus on both new and existing buildings. For the current study only the focus of these programmes on existing buildings was considered.

** For an extensive description of the various voluntary programmes studied see: www.EnviroVoluntarism.info
| Interviewee background | Government | Non-government |
|------------------------|------------|----------------|
| Policy maker           | 9 (4A/4N/1U)* |                |
| Administrator          | 29 (22A/4N/3U) | 17 (12A/3N/12U) |
| Architect, engineer,   |             | 13 (7A/6N)     |
| advisor                |             |                |
| Contractor, developer  |             | 7 (3A/4N)      |
| Property owner         |             | 11 (4A/3N)     |
| Other                  |             | 8 (3A/3N/2U)   |
| **Total**              | **39 (26A/8N/4U)** | **62 (29A/19N/14U)** |

*Abbreviations: A=Australia; N=Netherlands; U=United States
Table 3 – Data indicating qualitative differences in observations (NOTE: NO COLOUR FOR PRINT VERSION; COLOUR FOR DIGITAL VERSION)

| Case* | Country  | Type** | Condition*** | Outcome** |
|-------|----------|--------|--------------|-----------|
|       |          |        | FI | Nm | Le | Pc | En | Gi | Gp | O1 | O2 |
| 4     | Australia| BARE   | +  | ++ | -- | -- | -- | +  | -- | +  | -  |
| 1a    | Australia| BAP    | ++ | +  | ++ | +  | ++ | -  | +  | ++ | -  |
| 1b    | Australia| BAP    | ++ | +  | ++ | ++ | ++ | -- | +  | ++ | -  |
| 63    | Australia| BAP    | +  | +  | -- | -- | -- | -- | -- | -- | -- |
| 9     | Australia| CONE   | ++ | -  | ++ | +  | ++ | -  | +  | ++ | -  |
| 6     | Australia| INFI   | ++ | -  | +  | -  | ++ | +  | +  | -- | -  |
| 64    | Australia| INFI   | ++ | -  | -  | +  | ++ | -  | -- | -- | -- |
| 21    | Netherland| BARE  | +  | -  | -- | -- | -- | +  | -- | -- | -- |
| 23    | Netherland| BARE  | ++ | -  | -  | -  | -  | -  | +  | ++ | -  |
| 27a   | Netherland| BAP   | ++ | +  | ++ | +  | -- | +  | +  | -- | -- |
| 27b   | Netherland| BAP   | ++ | +  | ++ | ++ | +  | -  | +  | +  | -- |
| 25    | Netherland| CONE  | -  | +  | -- | +  | ++ | +  | ++ | +  |
| 26    | Netherland| CONE  | -  | +  | +  | -  | -  | +  | ++ | -  |
| 24    | Netherland| INFI  | +  | -  | -  | +  | ++ | -  | -  | -  | -  |
| 49    | US       | BARE   | ++ | +  | +  | +  | -- | +  | -- | -  | -  |
| 54    | US       | BARE   | ++ | +  | +  | -  | +  | -- | +  | ++ | +  |
| 55a   | US       | BAP    | ++ | +  | ++ | +  | -- | +  | ++ | -  |
| 55b   | US       | BAP    | ++ | +  | ++ | ++ | +  | -- | +  | ++ | -- |
| 50    | US       | CONE   | +  | -  | -- | -  | +  | -- | -- | -  | -  |
| 14    | US       | INFI   | ++ | -  | -- | +  | ++ | -  | -  | -  | -  |
| 22    | US       | INFI   | ++ | +  | +  | +  | -- | -- | +  | -  | -  |
| 42    | US       | INFI   | +  | -  | -- | +  | ++ | -  | -  | -- | -  |
| 46    | US       | INFI   | +  | -  | -- | +  | ++ | -  | -- | ++ | +  |

* Cases are given numbers to maintain anonymity as requested by some interviewees. Please note the ‘a’ and ‘b’ cases refer to specific arrangements that allow their participants to meet either high or moderate participation criteria.

** Approach to goal achievement: BARE = regulatory barrier relief; BAP = building assessment programme; CONE = collaborative network; INFI = innovative form of financing.

*** Conditions and outcomes as per section 3.4.

Notes: ++ = score (e.g., the arrangement has attracted at least the expected number of participants); + = score closer to “++” than to “--” (e.g., the arrangement has attracted a substantial number of participants, but not the expected number); - = score closer to “--” than to “++” (e.g., the arrangement has attracted a marginal number of participants, but this number is far from meeting the expected number); -- = minimum score (e.g., the arrangement has not attracted any or only a few participants).
Table 4 – Data support (or not) for expectations

| Expectation derived from literature | Outcome O2 (building retrofits) |
|-------------------------------------|--------------------------------|
|                                     | Consist. | Cover. |
| $F_g$ High financial gain, high outcomes | 0.95     | 0.35   |
| $N_m$ High non-monetary gain, high outcomes | 0.73     | 0.42   |
| $L_e$ High recognition and rewarding of leadership, high outcomes | 0.58     | 0.37   |
| $P_c$ Low participation criteria, high outcomes | 0.73     | 0.61   |
| $E_n$ Low enforcement, high outcomes | 0.73     | 0.50   |
| $G_i$ High government involvement, high outcomes | 0.47     | 0.37   |
| $G_p$ High government participation, high outcomes | 0.82     | 0.48   |

*Note:* This table presents the results of a QCA analysis aiming to understand if any of the conditions studied is individually necessary for causing the outcome building retrofits. Consistency (Consist.) scores of above 0.90 may indicate a necessary condition. The coverage (Cover.) scores help to distinguish among relevant and trivial necessary conditions.
### Table 5 – Configurations for achieving substantial to high numbers of building retrofits

| Causal Condition                  | Solution |
|-----------------------------------|----------|
|                                   | 1       | 2       | 3       |
| **Participation motivations**     |  ●      | ●       |         |
| Financial gain for participants   |         |         |         |
| Non-monetary gain for participants|         |         | ●       |
| Showcasing leadership             |         | ●       |         |
| **Stringency of the arrangement**|         |         |         |
| Participation criteria            | ø       | ø       | ●       |
| Enforcement of these criteria     |         |         | ●       |
| **Role of government in an arrangement** |         |         |         |
| Government involvement            | ø       | ●       |         |
| Government participation/launching customer | ø       |         |         |
| Raw coverage                      | 0.47    | 0.52    | 0.26    |
| Unique coverage                   | 0.05    | 0.16    | 0.05    |
| Consistency                       | 0.75    | 0.77    | 0.72    |
| Overall solution consistency      | 0.79    |         |         |
| Overall solution coverage         | 0.65    |         |         |

**Note:** This table presents the results of a set-theoretic analysis for the outcome O2 (substantial to high number of building retrofits). The analysis procedure has been well-documented elsewhere and is followed according to established QCA practice (Ragin, 2008; Ragin et al., 2006; Rihoux & Ragin, 2009; Schneider & Wagemann, 2012) – see further the supplementary appendix. The three ‘solutions’ are the logically reduced empirical observations of conditions that are sufficient for causing the outcome under scrutiny.
Table 6 – Configurations for not achieving any or only a marginal number of building retrofits

| Causal Condition | 4 | 5 | 6 | 7 | 8 |
|------------------|---|---|---|---|---|
| **Participation motivations** |   |   |   |   |   |
| Financial gain for participants | ø |   |   | ● |   |
| Non-monetary gain for participants | ø | ø |   |   |   |
| Showcasing leadership | ø | ø | ø |   |   |
| **Stringency of the arrangement** |   |   |   |   |   |
| Participation criteria | ● | ● |   | ● |   |
| Enforcement of these criteria | ● | ● | ● | ø |   |
| **Role of government in an arrangement** |   |   |   |   |   |
| Government involvement | ø | ø | ø | ● |   |
| Government participation/launching customer | ø | ø | ø | ø |   |
| **Raw coverage** | 0.13 | 0.45 | 0.18 | 0.36 | 0.31 |
| **Unique coverage** | 0.02 | 0.29 | 0.02 | 0.07 | 0.07 |
| **Consistency** | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| **Overall solution consistency** |   |   |   |   |   |
| **Overall solution coverage** | 1.00 |   |   | 0.82 |   |

Note: This table presents the results of a set-theoretic analysis for the outcome O2 (only a marginal number of building retrofits). The analysis procedure has been well-documented elsewhere and is followed according to established QCA practice (Ragin, 2008; Ragin et al., 2006; Rihoux & Ragin, 2009; Schneider & Wagemann, 2012) – see further the supplementary appendix. The five ‘solutions’ are the logically reduced empirical observations of conditions that are sufficient for causing the outcome under scrutiny.


| Interviewee* | Quote |
|--------------|-------|
| Int.82       | It’s pampering. Yes, of course, if you give people twelve hundred euros [through a voluntary programme] they will do something. |
| Int.45       | I think a lot of [retrofitting buildings] has to be driven by tenants or consumers. We [a major Australian property owner] can provide office buildings and houses [with high ratings in a voluntary program], but if people are not willing to pay for it then essentially we are working for a loss. |
| Int.50       | We [an administrator of a voluntary programme] found that the first thing people ask is “Well, what’s in it for me?” |

* In line with the custom of qualitative social science research, interviewees provided their insights in confidence. As such I cannot provide their identities. To give the reader insight into the variance within the interviews I refer to them with a number (e.g., ‘Int.50’).
| Interviewee* | Quote |
|--------------|-------|
| Int.180      | There has been a rise of uptake of green building practice, but this is mostly in the top-end of town. The class A buildings. But if you look at the class B or C buildings it is very little. And that is a challenge to change. The same incentive is not there. These are the smaller buildings. The older the building. They come with particular challenges. |
| Int.47       | However, I should note that we are talking about the top end of town here [where voluntary programmes achieve positive outcomes]. But there is another level [referring to households] where the consumers do not currently see the benefit of [sustainable buildings] and they don’t want to pay for it. And even if they do see the benefit, they probably are not willing to pay a premium for it. This is the next major challenge. |
| Int.31       | I think that in the commercial sector the long term benefits of sustainability features are realised much sooner [than in the residential sector]. Therefore it is much easier to promote. |
References
Ackerman, F. (1997). *Why do we recycle?* Washington: Island Press.
Baron, D. P., & Diermeier, D. (2007). Strategic Activism and Nonmarket Strategy. *Journal of Economics and Management Strategy, 16*, 599-634.
Borck, J., & Coglianese, C. (2009). Voluntary Environmental Programs. *Annual Review of Environmental Resources, 34*, 305-324.
Boyd, S. (2013). Financing and Managing Energy Projects Through Revolving Loan Funds. *Sustainability, 6*(6), 345-352.
Bressers, H., De Bruijn, T., & Lulofs, K. (2009). Environmental negotiated agreement in the Netherlands. *Environmental Politics, 18*(1), 58-77.
Briscoe, F., & Safford, S. (2008). The Nixon-in-China Effect. *Administrative Science Quarterly, 53*(3), 460-491.
Cialdini, R. (2009). *Influence*. New York: William Morrow.
Cole, R., & Valdebenito, M. J. (2013). The importation of building environmental certification systems. *BRI, 41*(6), 662-676.
Croci, E. (2005). *The Handbook of Environmental Voluntary Agreements*. Dordrecht: Springer.
Dixon, T., Ennis-Reynolds, G., Roberts, C., & Sims, S. (2009). Is there a demand for sustainable offices? *Journal of Property Research, 26*(1), 61-85.
EC. (2013). *Financial support for energy efficiency in buildings. Report SWD(2013) 143 final*. Brussels: European Commission.
EIA. (2013). How much electricity does an American home use? Retrieved 11 January 2014, from [http://www.eia.gov/tools/faqs/faq.cfm?id=97&t=3](http://www.eia.gov/tools/faqs/faq.cfm?id=97&t=3)
Eichholtz, P., Kok, N., & Quigley, J. (2010). Doing Well by Doing Good? *American Economic Review, 100*, 2492–2509.
Erkens, D., & Van der Stede, W. (2013). *Strategy and Control*. University of Southern California. Los Angeles.
Fiss, P. (2011). Building Better Causal Theories. *Academy of Management Journal, 52*(2), 393-420.
Fowler, K. M., & Rauch, E. M. (2006). *Sustainable Building Rating Systems: Summary*. Richland: Pacific Northwest National Laboratory.
Galvin, R. (2014). Why German homeowners are reluctant to retrofit. *BRI, 42*(4), 398-408.
Gillingham, K., Harding, M., & Rapson, D. (2012). Split Incentives in Residential Energy Consumption. *Energy Journal, 33*(2), 37-62.
Givens, J., & Jorgenson, A. (2013). Individual environmental concern in the world polity. *Social Science Research, 42*(2), 418-431.
Gollagher, M., & Hartz-Karp, J. (2013). The Role of Deliberative Collaborative Governance in Achieving Sustainable Cities. *Sustainability, 5*(6), 2343-2366.
Gunningham, N. (2009). The New Collaborative Governance. *Journal of Law and Society, 36*(1), 145-166.
Hassler, U. (2009). Long-term building stock survival and intergenerational management. *BRI, 37*(5-6), 552-568.
Hassler, U., & Kohler, N. (2014). The ideal of resilient systems and questions for continuity. *BRI, 42*(2), 158-167.
Hoffmann, M. (2011). *Climate Governance at the Crossroads*. Oxford: Oxford University Press.
Hofman, P., & De Bruijn, T. (2010). The Emergence of Sustainable Innovation. In J. Sarking, J. Cordeiro & D. Vasquez Bruzt (Eds.), *Facilitating Sustainable Innovation Through Collaboration* (pp. 115-133). Amsterdam: Springer.
IEA. (2009). *World Energy Outlook 2009*. Paris: International Energy Agency/OECD.

IEA. (2012). *Tracking Clean Energy Progress*. Paris: International Energy Agency.

IEA. (2013). *Modernising Building Energy Codes*. Paris: United Nations Development Programme.

IPCC. (2014). *Climate Change 2014*. Cambridge: Cambridge University Press.

Kollmuss, A., & Agyman, J. (2002). Mind the Gap. *Environmental Education Research, 8*(3), 239-260.

Liu, F., Meyer, A., & Hogan, H. (2010). *Mainstreaming Building Energy Efficiency Codes in Developing Countries*. New York: The World Bank.

Managan, K., Layke, J., Monica, A., & Nesler, C. (2012). *Driving Transformation to Energy Efficient Buildings*. Washington: Institute for Building Efficiency.

Nash, J., & Revesz, R. L. (2007). Grandfathering and Environmental Regulation. *Northwestern University Law Review, 101*(4), 1677-1734.

NeJaime, D. (2009). When New Governance Fails. *Ohio State Law Journal, 70*(2), 323-399.

OECD. (2003). *Voluntary Approaches for Environmental Policy*. Paris: OECD.

OECD. (2013). *Greening Household Behavior*. Paris: OECD Publishing.

Osbaldiston, R., & Schott, J. P. (2012). Environmental Sustainability and Behavioral Science. *Environment and Behavior, 44*(2), 257-299.

Pivo, G. (2010). Owner-Tenant Engagement in Sustainable Property Investing. *The Journal of Sustainable Real Estate, 2*(1), 183-199.

Potoski, M., & Prakash, A. (2009). *Voluntary Programs*. Cambridge: MIT Press.

Potoski, M., & Prakash, A. (2013). Do Voluntary Programs Reduce Pollution? *Policy Studies Journal, 41*(2), 273-294.

Prakash, A., & Potoski, M. (2012). Voluntary environmental programs. *Journal of Policy Analysis and Management, 31*(1), 123-138.

Ragin, C. (2008). *Redesigning Social Inquiry*. Chicago: Chicago University Press.

Ragin, C., Rubinson, C., Schaefer, D., Anderson, S., Williams, E., & Giesel, H. (2006). *User’s Guide to Fuzzy-Set Qualitative Comparative Analysis*. Tucson: University of Arizona.

Rihoux, B. (2013). Qualitative Comparative Analysis (QCA), Anno 2013. *Swiss Political Science Review, 19*(2), 233-245.

Rihoux, B., & Ragin, C. (2009). *Configurational Comparative Analysis*. London: Sage.

Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public Administration, 80*(1), 1-22.

Schneider, C., & Wagemann, C. (2012). *Set-Theoretic Methods for the Social Sciences*. Cambridge: Cambridge University Press.

Short, J., & Toffel, M. W. (2010). Making Self-Regulation More Than Merely Symbolic. *Administrative Science Quarterly, 55*(2), 361-396.

Sunikka-Blank, M., & Galvin, R. (2012). Introducing the prebound effect. *BRI, 40*(3), 260-273.

Thøgersen, J. (2010). Country Differences in Sustainable Consumption. *Journal of Macromarketing, 30*(2), 171-185.

UNDP. (2013). *Human Development Report 2013*. New York: United Nations Development Programme.

Van der Heijden, J. (2012). Voluntary Environmental Governance Arrangements. *Environmental Policies, 21*(3), 486-509.

Van der Heijden, J. (2014). *Governance for Urban Sustainability and Resilience*. Cheltenham: Edward Elgar.
Venesson, P. (2008). Case studies and process tracing. In D. Della Porta & M. Keating (Eds.), *Approaches and Methodologies in the Social Sciences* (pp. 223-239). Cambridge: Cambridge University Press.

Vinagre Diaz, J. J., Wilby, M. R., & Belén Rodríguez González, A. (2013). Setting up GHG-based energy efficiency targets in buildings. *Energy Policy*, 59(C), 633–642.

Yudelson, J., & Meyer, U. (2013). *The World’s Greenest Buildings*. Abingdon: Routledge.
Appendix: Applying fuzzy set qualitative comparative analysis (fsQCA) in this study

QCA was introduced by Charles Ragin, a social scientist, as a middle path between quantitative and qualitative social research (Ragin, 1987). QCA is grounded in set theory, a branch of mathematical logic that allows studying in detail how causal conditions contribute to a particular outcome. QCA has since the mid-1990s quickly evolved as an accepted research practice for the type of study presented in this article, and has been applied in hundreds of studies in the policy sciences in particular and the social sciences more generally (Rihoux, 2013; Rihoux & Ragin, 2009). See for instance a recent special issue of Policy and Society on 'Innovative methods for policy analysis: QCA and fuzzy sets' (Policy and Society, 2013).

Particularly the introduction of fsQCA appears to have spurred the use of the method by a range of scholars from various backgrounds studying issues such as governance networks within a country (Verweij, Klijn, Edelenbos, & Van Buuren, 2013), job security regulations in Western democracies (Emmenegger, 2011), or organisational configurations (Fiss, 2011). In a recent article, Barbara Vis (2012) has compared fsQCA with Regression Analysis by applying the two data analysis techniques to a single dataset. She finds that fsQCA is better capable to understand and make visible complex relations than regression analysis (for a comparable argument, see Warren, Wistow, & Bambra, 2013).

The fundamentals and background of the method are well-explained and documented in a series of strong textbooks (Goertz & Mahony, 2012; Ragin, 2008; Rihoux & Ragin, 2009; Schneider & Wagemann, 2012). These handbooks are good further references for those unfamiliar with the foundations of the method (which I will not dwell on here). These handbooks provide guidelines for good fsQCA practice (Ragin, 2008, see in particular the 'practical appendices'; Rihoux & Ragin, 2009, particularly Chapter 5; Schneider & Wagemann, 2012, Chapter 11), which I have followed closely in carrying out the analyses discussed in the article. One of the key points for good fsQCA practice is for the researcher to provide as much transparency into the analysis as possible. This is what is seek to do through this supplementary appendix.

In this appendix I give account of the various steps taken and decisions made in the fsQCA analysis as presented in the article. Where necessary I provide additional (raw) data so that the interested reader can repeat my analysis. In giving account I follow the ‘flowchart’ of Jerry Mendel and Mohammad Korjani (2013) who, supported by Charles Ragin, have mathematically summarized fsQCA as a collection of 13 steps. I do however take the liberty to use the jargon from the handbooks (as compared to the mathematical jargon introduced by Mendel and Korjani) to specify the steps and collide them into 10 steps.

In addition to Mendel and Korjani’s steps of how the fsQCA analysis is carried out it is, of course, of importance to motivate why fsQCA was chosen in the first place. Whilst researchers often support their choice for fsQCA with a practical motivation (i.e., they have a medium number of cases that likely allows for systematic cross-case analysis, but not for sophisticated statistical analysis), ideally fsQCA is chosen for a theoretical motivation (Schneider & Wagemann, 2012). I have added this step of motivating the choice for fsQCA, and will start with it in what follows.

Step 1: Why an fsQCA analysis?
Earlier empirical studies have found that, for instance, voluntary programmes with similar designs such as pay-per-plastic-bag fees (Ackerman, 1997), organic food labelling (Thøgersen, 2010), building assessment classification and certification (Fowler & Rauch, 2006), and revolving loan funds (Boyd, 2013) show different outcomes depending on how
their design conditions interact with contextual conditions (e.g., existing legislation, economic circumstances; Borck & Coglianese, 2009). Even more, some studies indicate that a single design (e.g., building assessment classification and certification) implemented in a number of similar contexts (e.g., the United States, Australia, the United Kingdom) may nevertheless result in different outcomes due to the role of governmental actors in these arrangements (Fowler & Rauch, 2006).

This all indicates that the outcomes of voluntary programmes are likely caused by different interacting conditions (i.e., conjunctural causation), that different (configurations of interacting) conditions may cause a similar outcome (i.e., equifinality), and that the presence of a (configuration of interacting) condition(s) in the causal role of the outcome is of limited help in explaining the inverse situation (that is, the causal role of the absence of the condition in the non-occurrence of the outcome; i.e., asymmetry).

QCA is chosen as a data analysis methodology because it allows for ‘unraveling causally complex patterns in terms of equifinality, conjunctural causation, and asymmetry’ (Schneider & Wagemann, 2012, 8). QCA differs from other data analysis methods in its focus. ‘The key issue [for QCA] is not which variable is the strongest (i.e., has the biggest net effect) but how different conditions combine and whether there is only one combination or several different combinations of conditions (causal recipes) of generating the same outcome’ (Ragin, 2008, 114). QCA helps to trace patterns of association between these conditions in a highly systemised manner and allows for systematic comparisons between empirical observations (i.e., cross-case), whilst allowing for in-depth within-case understanding of the individual observations (Rihoux & Ragin, 2009).

I have chosen fuzzy set QCA (fsQCA) as it allows for giving a rather precise insight in the qualitative difference in my empirical data – i.e., the degree of presence or absence of a condition or the outcome in the cases under analysis. I will explain this particular issue to more depth under step under step 4.

**Step 2:** Selection of outcome of interest and cases to study
In the article I explain the outcomes of interest. The operationalization of these outcomes is further explained below. The selection of cases (real world examples of innovative governance voluntary programmes) is also explained in the article. In short, I have selected 20 cases of voluntary programmes from a larger study (Van der Heijden, forthcoming, 2014).

**Step 3:** Select k causal conditions
In the article I discuss a set of seven design conditions that the current literature considers most relevant in influencing the performance of voluntary programmes. The operationalization of these conditions is explained below.

**Step 4:** Calibration of set-membership scores for outcomes and conditions
The strength of fsQCA as compared to other forms of QCA is that it allows for giving a rather precise insight in the qualitative difference in the units of observation. In other words, it allows distinguishing among different qualitative categories of these observations and compare sets of observations of a particular category with sets of observations of other categories.

To illustrate, imagine that you have to classify the greenness of 10 paintings for a particular analysis. You could, of course, measure the percentage of green in the paintings and rank them accordingly, but how to deal with different shades of green? Or, is a fully green 2”x2” painting greener than a 3’2”x3’2” that has ‘only’ 50% of its canvas painted green? And
what about a painting that is predominantly yellow and blue, the prime colours that together make green?

Before you start this seemingly easy task, you will have to come up with at least two categories for ranking the paintings: one for the paintings that meet a certain understanding of green, and the others that do not. Let’s assume that you decide that in order to be considered ‘green’ half of a painting needs to be painted a shade of green. This categorisation results in seven paintings meeting the condition green, and three paintings that do not. Upon second inspection you find that two of those seven in the “in” category are very green (say, 80% or more); and one of the three paintings in the “out” category has no green at all, leaving two paintings somewhat green. In analysing the paintings it may be of interest to use these qualitative differences in the original “in” and “out” categories. You therefore decide to group those two paintings from the original “in” category as having “full-membership” in the condition green and the five remaining paintings as being “more in than out” of the condition green. The two somewhat green paintings in the original “out” category could be further grouped as “more out than in” the condition green, leaving the last painting in the category “full non-membership” in the condition green.

This is precisely what calibration of data in fsQCA implies. It asks the researcher to carefully distinguish the various qualitative categories of their observations according to their qualitative differences and carefully assign their data to these categories. Good fsQCA practice requires the researcher to be clear about this calibration. Particularly to explain the two extremes of the observed data (i.e., maximum and minimum fit in a category), and the crossover point of the data (i.e., in what stage is the data considered to have maximum ambiguity; that is, when is it as much in as out?) (Ragin, 2008; Rihoux & Ragin, 2009; Schneider & Wagemann, 2012). I have calibrated my data using a four category qualitative scale as represented in table A1.

Table A1 – Verbal description of membership scores of the data in qualitative categories

| The observation is...                                                                 | Qualitative symbol | Fuzzy set value |
|--------------------------------------------------------------------------------------|--------------------|-----------------|
| Full membership (i.e., in the highest stage observed)                                | ++                 | 1.00            |
| More in than out                                                                     | +                  | 0.67            |
| More out than in                                                                     | -                  | 0.33            |
| Full non-membership (i.e., in the lowest stage observed)                              | --                 | 0.00            |

For the various outcomes and conditions the extremes and crossover points in the data are set as follows:

1. Observed outcome:
   - Participants (O1). I have operationalised this outcome by considering the participants a voluntary programme has attracted as compared with stated ambitions (in documentation or expressed in interviews). Full membership represents that stated ambitions are met; full non-membership represents not having attracted any or only a few participants; and, the crossover point is set at not meeting half the stated ambitions in terms of attracting participants.
   - Building retrofits (O2). I have operationalised this outcome by considering the number of building retrofits achieved under a voluntary programme as compared with stated
ambitions (in documentation or expressed in interviews). Full membership represents that stated ambitions are met; full non-membership represents not having achieved any building retrofits at all; and, the crossover point is set at not meeting half the stated ambitions in terms of building retrofits.

2. Conditions:
   o **Financial gain (Fg).** The qualitative categories for the direct financial gain (including cost savings) participants may get from joining an voluntary programme and retrofitting buildings are constructed by combining data on 'promised' gains (i.e., how prospective gains are marketed by the administrators of these voluntary programmes) and 'evidenced' gains (i.e., how realized gains are marketed by administrators and participants of these voluntary programmes). Full membership represents a marketed high certainty of achieving substantial financial gains when participating based on evidence. More in that out membership represents a marketed promised certainty of gains supplemented with evidence. More out than in represents a marketed promise of gains when participating. Full non-membership represents a full absence of a marketing of gains. The crossover point is the marketing of promised high certainty of gains but without evidence to support this promise.
   o **Non-monetary gain (Nm).** I followed the above line of reasoning. Thus, full membership represents a marketed high certainty of achieving substantial non-monetary gain when participating based on evidence; full non-membership represents a full absence of a marketing of non-monetary gains; and, the crossover point is the marketing of promised high certainty of non-monetary gains but without evidence to support this promise.
   o **Showcasing leadership (Le).** To construct a fuzzy set for this condition I considered how administrators of voluntary programmes reward and market leadership. Full membership represents a focus on national or global leadership combined with marketing of leading practice or awarding of leading practice through, for instance, yearly awarding ceremonies. More in than out represents a focus on regional or local leadership combined with marketing or awarding of such leadership. More out than in represents a focus on leadership in the marketing of a voluntary programme, but an absence of marketing or rewarding actual leadership by participants. Fully non-membership represents a full absence of a focus on leadership in the marketing of a voluntary programme. The crossover point of this condition is the marketing of best-practices as opposed to local, national or international leadership.
   o **Participation criteria (Pc).** I followed the above line of reasoning. Thus, full membership represents that participants are required to perform significantly beyond the requirements of public law and regulation (e.g., to achieve double the statutory requirement, or to show high level performance in an area that is not yet addressed through statutory regulation). More in than out represents that participants are required to perform well beyond the requirements of public law and regulation (e.g., to achieve more than the statutory requirement, or to show unspecified performance in an area that is not yet addressed through statutory regulation). More out than in represents that participants are required to perform just beyond the requirements of public law and regulation. Full non-membership represents a full absence of criteria. The crossover point of this condition is set at criteria that only require performance that is marginally better than what is required by statutory regulation.

---

11 Please note, a voluntary programme may for instance seek to ease or stimulate compliance with statutory regulation.
• *Enforcement criteria (En).* The qualitative categories reflect the strictness of enforcement in terms of who enforces, how enforcement is carried out, and what evidence results from enforcement. Full-membership represents strict enforcement; for instance third-party enforcers, a documented enforcement process, and the awarding of a certificate at the end of the process. More in than out represents medium enforcement; for instance, administrator enforcement and documented proof of compliance at the end of the process. More out than in represents a weak enforcement process; for instance, participant self-enforcement. Full non-membership represents the absence of enforcement. The crossover point of this condition is set at administrator enforcement without documented proof of compliance at the end of the enforcement process.

• *Government involvement (Gi).* To construct a fuzzy set for this condition I have considered how governments are involved in the voluntary programmes. Full membership represents sole governmental involvement in initiating and administrating a voluntary programme. More in than out represents sole governmental involvement in initiating a voluntary programme. More out than in represents equal involvement in the initiating of voluntary programmes of governmental and non-governmental actors. Full non-membership represents the absence of government involvement. The crossover point of this condition is set at dominance of governmental involvement in this role.

• *Government participation (GP).* Full membership represent high activity; for instance, government as dominant participants or customers of an voluntary programme as a result of mandatory participation or procurement. More in than out represents medium activity; for instance, mandatory participation or procurement criteria (but no government dominance in an voluntary programme’s participants or customers). More out than in represents low activity; for instance, preferred participation or procurement. Full non-membership represents no government role. The crossover point is set at specified requirements for governments to participate in, or require their suppliers to participate in specific voluntary programmes.

**Step 5: Create a raw data matrix**

Now that the various qualitative differences of the outcomes and conditions have been distinguished the data can be transformed into a raw data matrix. Table A2 provides the raw data used in the article.
Table A2 – Raw data matrix

| Case* | Condition** | Outcome** |
|-------|-------------|-----------|
|       | Fi | Nm | Le | Pc | En | Gi | Gp | O1 | O2 |
| 1a    | 1.00 | 0.67 | 1.00 | 0.67 | 1.00 | 0.33 | 0.67 | 1.00 | 0.33 |
| 1b    | 1.00 | 0.67 | 1.00 | 1.00 | 1.00 | 0.33 | 0.67 | 1.00 | 0.00 |
| 4     | 0.67 | 1.00 | 0.00 | 0.33 | 0.00 | 0.67 | 0.00 | 0.67 | 0.33 |
| 6     | 1.00 | 0.33 | 0.33 | 0.67 | 0.33 | 1.00 | 0.67 | 0.67 | 0.00 |
| 9     | 1.00 | 0.33 | 1.00 | 1.00 | 0.67 | 1.00 | 0.33 | 0.67 | 1.00 |
| 14    | 1.00 | 0.00 | 0.00 | 0.67 | 0.67 | 1.00 | 0.33 | 0.33 | 0.33 |
| 21    | 0.67 | 0.33 | 0.00 | 0.00 | 0.00 | 0.67 | 0.33 | 0.33 | 0.00 |
| 22    | 1.00 | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 | 0.00 | 0.00 | 0.67 |
| 23    | 1.00 | 0.33 | 0.33 | 0.33 | 0.33 | 0.33 | 0.33 | 1.00 | 0.67 |
| 24    | 0.67 | 0.33 | 0.33 | 0.67 | 0.67 | 1.00 | 0.33 | 0.33 | 0.33 |
| 25    | 0.33 | 0.67 | 0.00 | 0.67 | 0.67 | 0.67 | 1.00 | 0.67 | 1.00 |
| 26    | 0.33 | 0.67 | 0.67 | 0.67 | 0.33 | 0.33 | 0.67 | 1.00 | 0.33 |
| 27a   | 1.00 | 0.67 | 1.00 | 0.67 | 0.67 | 0.33 | 0.67 | 1.00 | 0.33 |
| 27b   | 1.00 | 0.67 | 1.00 | 1.00 | 0.67 | 0.33 | 0.67 | 0.67 | 0.00 |
| 42    | 0.67 | 0.33 | 0.00 | 0.67 | 0.67 | 1.00 | 0.33 | 0.33 | 0.00 |
| 46    | 0.67 | 0.33 | 0.00 | 0.33 | 0.67 | 0.67 | 0.00 | 1.00 | 0.67 |
| 49    | 1.00 | 0.67 | 0.67 | 0.67 | 0.00 | 0.67 | 0.00 | 0.33 | 0.33 |
| 50    | 0.67 | 0.33 | 0.00 | 0.33 | 0.67 | 0.33 | 0.00 | 0.33 | 0.33 |
| 54    | 1.00 | 0.67 | 0.67 | 0.67 | 0.33 | 0.67 | 0.00 | 1.00 | 0.67 |
| 55a   | 1.00 | 0.67 | 1.00 | 0.67 | 0.67 | 0.00 | 0.67 | 1.00 | 0.33 |
| 55b   | 1.00 | 0.67 | 1.00 | 1.00 | 0.67 | 0.00 | 0.67 | 1.00 | 0.00 |
| 63    | 0.67 | 0.67 | 0.00 | 0.67 | 0.33 | 0.00 | 0.00 | 0.33 | 0.00 |
| 64    | 1.00 | 0.33 | 0.33 | 0.67 | 0.33 | 1.00 | 0.67 | 0.33 | 0.00 |

* Cases are given numbers to maintain anonymity as requested by some interviewees. Please note the ‘a’ and ‘b’ cases refer to specific arrangements that allow their participants to meet either high or moderate participation criteria.

** Conditions and outcomes as per Section 2.4 in the article.

Notes: 1.00 = maximum score (e.g., the arrangement has attracted at least the expected number of participants); 0.67 = score closer to “1.00” than to “0.00” (e.g., the arrangement has attracted a substantial number of participants, but not the expected number); 0.33 = score closer to “0.00” than to “1.00” (e.g., the arrangement has attracted a marginal number of participants, but this number far from meeting the expected number); 0.00 = minimum score (e.g., the arrangement has not attracted any or only a few participants)

Step 6: Analysis of necessary conditions

Following established fsQCA practice the data is first analysed for necessary conditions before exposing it to more complex analysis to identify (configurations of) sufficient conditions (Rihoux & Ragin, 2009, Chapter 5, box 8.1; Schneider & Wagemann, 2012, Chapter 11).

For a condition to be necessary to cause the outcome the fuzzy-set membership scores of the outcome need to be a perfect subset of the membership scores of the condition.

To give an illustration, in order for it to be true that one can practice as an architect in the Netherlands (outcome) it is necessary that one is registered with the Netherlands Architects...
Registrar (condition). In other words, the set of architects in the Netherlands (outcome) is a subset of individuals registered with the Netherlands Architects Registrar (condition). The set of individuals registered with the Netherlands Architects Registrar is however (much) larger than the set of architects because it also includes landscape architects, city planners and interior designers.

To gain an insight as to whether any of the distinct conditions (see step 3) is necessary for causing the outcome (see step 3) I have plotted a series of fuzzy set XY plots using the computer program FS/QCA version 2.5 (cf., Schneider & Wagemann, 2012, Chapter 5 and Chapter 9). Table 4 in the article presents the results.

In studying necessity two issues are of importance: consistency and coverage. Consistency indicates how strongly the condition relates to the outcome. In other words, if a hypothesised relation between a condition and an outcome is not consistent (where the advisory cut-off point of consistency is a score of 0.90), the hypothesised relation cannot be supported by the data as being necessary (Rihoux & Ragin, 2009, 45). Table 4 in the article indicates that only the condition ‘financial gain’ passes the consistency test. However, the low coverage score of 0.35 indicates that this is likely a trivial necessary condition in achieving this outcome (Schneider & Wagemann, 2012, 232-237).

Coverage indicates how relevant the condition is for causing the outcome. Coverage is only assessed for conditions that meet the consistency test. Here it is important to distinguish between relevant and trivial necessary conditions. In other words, if a consistent relation only covers a small number of cases (i.e., if it has a low coverage score such as the condition ‘financial gain’) it can be considered to be trivial in causing the outcome (further, Schneider & Wagemann, 2012, Chapter 9). Another way to distinguish between relevant and necessary conditions is to assess whether or not the data is skewed towards conditions that have high scores for both the condition and the outcome. This suggests that such conditions may pass the test for both necessity and sufficiency, and is likely a trivial necessary condition in achieving this outcome (Schneider & Wagemann, 2012, 232-237).

**Step 7: Analysis of sufficient conditions (1): create a truth table**

Having studied the data for necessary conditions (but having found none) the next step is to study the data for sufficient conditions.

For a condition or for a configuration of conditions to be sufficient for causing the outcome the fuzzy-set membership scores of the condition or the configuration of conditions need to be a perfect subset of the membership scores of the outcome. To give an illustration, whilst being registered with the Netherlands Architects Registrar is a necessary condition for one to practice as an architect in the Netherlands this is however not a sufficient condition. After all, one also needs an office, the relevant design software, and so on, to be able to practice as an architect (i.e., other necessary conditions). In order to get the registration (outcome) it is however sufficient that one holds a Masters’ Degree in Architecture from Delft University of Technology (condition), which the Registrar accepts as meeting the requirements for registration. This Degree is however not necessary for registration since the Registrar accepts degrees from other educational facilities as well (i.e., other sufficient conditions).

The analysis of (configurations of) sufficient conditions for the outcomes under scrutiny follows three sub-steps. The first sub-step is to create a truth table. Established QCA practice requires to present this truth table because it is the basis of the following analysis. Table A3 provides the truth table for the analysis of sufficient conditions for the outcome under scrutiny (‘O2’). The truth table is created using FS/QCA software (version 2.5).
Table A3 - Truth table for the analysis of sufficient conditions for outcome O2

| Row | Conditions | Number (case) | Outcome O2 | Consistency Raw |
|-----|------------|---------------|------------|-----------------|
| 1   | 1 0 0 0 1 1 0 | 1 1 | 1 | 0.82 |
| 2   | 0 1 0 1 1 1 1 | 1 1 | 1 | 0.80 |
| 3   | 1 0 0 0 0 0 0 | 1 1 | 1 | 0.78 |
| 4   | 1 1 0 1 0 0 0 | 1 0 | 0 | 0.76 |
| 5   | 1 0 0 1 0 0 1 | 0 1 | 1 | 0.75 |
| 6   | 1 1 1 1 0 1 0 | 1 1 | 1 | 0.67 |
| 7   | 1 0 0 0 0 1 0 | 1 0 | 0 | 0.66 |
| 8   | 1 1 0 0 0 1 0 | 1 0 | 0 | 0.66 |
| 9   | 1 0 1 1 1 1 0 | 1 1 | 1 | 0.64 |
| 10  | 1 0 0 1 0 1 1 | 1 1 | 1 | 0.62 |
| 11  | 1 0 0 1 1 1 0 | 3 0 | 0 | 0.61 |
| 12  | 1 1 1 1 1 0 0 | 1 0 | 0 | 0.54 |
| 13  | 1 1 1 1 0 0 1 | 1 1 | 1 | 0.49 |
| 14  | 1 1 1 1 0 1 0 | 6 0 | 0 | 0.28 |
| 15-128 | Logical remainders | | | |

Note: The table and symbols are explained in the text below. Abbreviations as per table A2 and the article.

The truth table is a data matrix with $2^k$ rows that represents all possible configurations of conditions that are logically possible. Note, the truth table reports data using the cross-over points set – i.e., ‘1’ indicates more in than out of the set (including full membership), and ‘0’ indicates more out than in the set (including full non-membership). Thus, with the seven conditions here the number of logically possible configurations is 128 (i.e., $2^7$). The empirical observations are included in this table. As the truth table indicates, out of 128 logically possible configurations 14 were empirically observed (rows 1 to 14).

The different rows can be understood as ideal types (Schneider & Wagemann, 2012, Chapter 7). The number column (‘No.’) indicates how many cases fit best in this ideal type (i.e., when a case has a membership in the configuration of the fuzzy-sets for the conditions of at least 0.5). The row ‘outcome’ indicates whether for a configuration of conditions the outcome was observed or not (a ‘1’ indicates it is, a ‘0’ indicates it was not).

Because some observations of configurations of conditions may be observed in different cases, some rows in the truth table may refer to many cases (e.g., row 14) whilst other rows refer to only a few or just one case (e.g., rows 1 and 11). It is normal that the truth table also contains rows of possible combinations, but without empirical observations (i.e., rows 15 - 128).

In the second sub-step the truth table is logically minimized based on two conditions. First, the researcher sets a threshold for ‘logical remainders’. Logical remainders are those configurations of conditions that ‘lack enough empirical evidence to be subjected to a test of sufficiency’ (Schneider & Wagemann, 2012, 152). It depends on the size of the research project (i.e., the number of cases included) what is to be considered as ‘enough empirical
evidence’. Most often a threshold of one observation (thus at least one case) is used, but for larger numbers of cases a higher threshold can be applied (Ragin, 2008; Schneider & Wagemann, 2012). Following this practice I have decided a threshold of at least one observation.

Second, the researcher has to set a ‘consistency threshold for distinguishing [configurations of conditions] that are subsets of the outcome from those that are not’ (Ragin, 2008, 143). In other words, how well do the configurations of conditions fit the outcome? This is what the ‘raw consistency’ score in the truth table indicates. As discussed under step 6, the higher the score the better the fit. Ragin (2008) advises a consistency score of at least 0.75, which I have followed. Please note, the consistency score of row 5 before rounding was < 0.75.

This resulted in a minimization in which 4 cases met the observation threshold; and two cases met the consistency threshold for the outcome under scrutiny. In FS/QCA cases that met the consistency threshold were labelled ‘1’ in the outcome column in the and those that did not were labelled ‘0’ (cf., Ragin, 2008, 144).

Step 8: Analysis of sufficient conditions (2): dealing with logical remainders and choice of solution term

Having carried out this minimization of the truth table a standard analysis can be run in FS/QCA (the third sub-step). This standard analysis is best understood as the identification of ‘the combinations of attributes [i.e., configurations of necessary conditions] associated with the outcome of interest using Boolean algebra and algorithms that allow logical reduction of numerous, complex causal [configurations of] conditions into a reduced set of configurations that lead to the outcome’ (Fiss, 2011, 402). Normally a standard analysis results in a solution that consists of a number of ‘paths’ or ‘solutions’ (combinations of sufficient conditions) that lead to the outcome.

A simple example may explain what is going on in this analysis: imagine two painters painting a fully green painting. You are interested to see what colour green they are using. In inspecting their paint boxes you find that Painter A has yellow, blue and red paint; and Painter B has yellow and blue paint but not red. How then have they come to green?

First, they must have mixed their paints as blue, yellow, or red cannot suddenly become green (or at least, let us assume that for the sake of the example). Second, they have mixed at least two paints. But what mixes have they used?

Painter A could have mixed ‘blue’ and ‘yellow’ and ‘red’, whilst Painter B could have mixed ‘blue’ and ‘yellow’ but not ‘red’. It logically follows that ‘red’ is not needed to make green. In other words, red can be eliminated as a condition needed for the outcome green, which leaves the combination ‘blue and yellow’ as the configuration of paints that causes green. This is in a nutshell what the logical reduction of causal configurations of conditions implies (more sophisticated explanations are found in the handbooks by Goertz & Mahony, 2012; Ragin, 2008; Rihoux & Ragin, 2009; Schneider & Wagemann, 2012).

The standard analysis produces three types of logically reduced configurations of conditions that are sufficient for the outcome under scrutiny: a complex solution, an intermediate solution and a parsimonious solution.

The complex solution is exclusively based on the empirical information at hand. The complex solution can however be further simplified by using counterfactuals for the logical remainders. Distinction is made in ‘easy counterfactuals’ and ‘difficult counterfactuals’ (this is well explained by Fiss, 2011). Easy counterfactuals are based on the theoretical assumptions (or other substantive knowledge by the researcher), for this study these are the assumptions
identified in table A6 (above). Including easy counterfactuals in the standard analysis leads to the intermediate solution. Another illustration may be helpful here.

Assume again those two painters, but this time you cannot observe the yellow paint in Painter A’s paint box (say, after mixing green Painter A gave her yellow to Painter B who forgot to bring it). If theoretical or otherwise substantive knowledge exist that in order to make green paint yellow paint is needed then the counterfactual ‘yellow’ may be added to the analysis. The intermediate solution would then (again) indicate that in order to get ‘green’ ‘blue and yellow’ are needed, and that ‘red’ can be eliminated as a condition.

The parsimonious solution (i.e., the most simplified solution) results from using difficult counterfactuals. Applying difficult counterfactuals is the inverse of applying easy counterfactuals. That is, assumptions are made about the outcome of a configuration if the counterfactual condition is redundant. This is a more complicated (and risky) undertaking since expectations are often made on conditions being present, and not absent. Note however that a parsimonious solution may look ‘simpler’ than an intermediate or complex solution, but in fact gives less specific insight. Again an illustration may be helpful.

Assume that Painter A and Painter B are again painting a green painting, but of a much lighter shade than before. Inspecting their paintboxes you find that Painter A has blue, yellow, and white; and Painter B has only blue, and yellow. Based on theoretical knowledge you may assume that adding white to the mix of blue and yellow will result in a light shade of green. Adding this easy counterfactual results in the following intermediary solution: the configuration of blue, yellow and white paint is sufficient to paint a painting in a light shade of green.

But can you also assume that not adding white to the mix will result in a light shade of green? Say that you could (for instance, you could argue that Painter B has painted the painting so thin that the white base layer of the painting shines through, making it overall a lighter shade of green). Using this difficult counterfactual results in the more simplified parsimonious solution: the configuration of blue and yellow paint is sufficient to paint a painting in a light shade of green.

However, the parsimonious solution allows for a much larger set of outcomes than paintings in a light shade of green only. It includes the set of outcomes that includes paintings in any shade of green. This indicates that falsely made assumptions about difficult counterfactuals do however not give ‘false’ solutions, they just give more inclusive solutions. Though, parsimonious solutions may be ‘unrealistically simplistic’ (Ragin, 2008, 175).

I have used the expectations (based on the theory on voluntary programmes) as expressed in section 2.4 of the article for these minimization steps.

**Step 9: Presentation of results**

After carrying out the standard analysis results can be presented in various forms. Table 5 in the article is one of the accepted formats for doing so. In the text under table 5 in the article I explain how this table can be read.

**Step 10: Interpretation of results (and repeat the steps)**

Of course, an fsQCA analysis is but a means to an end and not an end in itself. Once the above analysis has been carried out, the findings should be interpreted in the light of the data obtained. This is what I do in the second half of the article.

Besides, I am not only interested in better understanding the causes of the outcome under scrutiny, but also the causes that have not resulted in this outcome. I have carried out
that analysis in Section 3.3 of the article, following the above steps. Table A4 provides the truth table for this analysis.

Table A4 - Truth table for the analysis of sufficient conditions for outcome ‘not O2’

| Row | Conditions | Number (case) | Outcome not O2 | Consistency Raw |
|-----|------------|---------------|----------------|-----------------|
| 1   | 1 1 1 1 1 0 | 1 0 0 | 1 | 1 | 1.00 |
| 2   | 1 1 1 0 0 0 | 1 0 0 | 1 | 1 | 1.00 |
| 3   | 1 1 1 0 1 0 | 0 1 0 | 1 | 1 | 1.00 |
| 4   | 1 0 0 1 0 0 | 0 0 0 | 1 | 1 | 1.00 |
| 5   | 1 0 0 0 0 0 | 1 0 0 | 1 | 1 | 1.00 |
| 6   | 1 0 0 0 0 0 | 1 0 0 | 1 | 1 | 1.00 |
| 7   | 0 1 1 1 1 0 | 0 0 0 | 1 | 1 | 1.00 |
| 8   | 1 0 0 1 1 0 | 0 0 0 | 3 | 1 | 1.00 |
| 9   | 1 1 1 1 1 1 | 0 1 0 | 0 | 1 | 1.00 |
| 10  | 1 0 0 0 0 0 | 1 0 1 | 0 | 0 | 0.91 |
| 11  | 1 0 0 0 0 0 | 0 0 0 | 1 | 0 | 0.89 |
| 12  | 1 0 1 1 1 0 | 0 1 0 | 0 | 0 | 0.82 |
| 13  | 0 1 1 1 1 0 | 1 1 0 | 0 | 0 | 0.80 |
| 14  | 1 1 1 1 0 0 | 0 1 0 | 1 | 0 | 0.78 |
| 15-128 | Logical remainders | | | | |

Note: The table and symbols are explained in the text below. Abbreviations as per table A2 and the article.

Note, Ragin (2008) advices a consistency score of at least 0.75, which I have followed in the analysis presented under step 7. However, when possible a higher consistency score can be used. It is then advised to look for big jumps in the consistency score column (Schneider & Wagemann, 2012). For this analysis a big jump is identified between row 9 and row 10 (i.e., from 1.00 to 0.91). I have therefore chosen a consistency cut off of 1.00 for this analysis.

The findings from the analysis that seeks to understand which configurations of conditions are related to not having achieved the outcome is discussed in section 4.3 in the article, and the related table 6 in the article.
References

Ackerman, F. (1997). Why do we recycle? Markets, Value and Public Policy. Washington: Island Press.

Borck, J., & Coglianese, C. (2009). Voluntary Environmental Programs: Assessing Their Effectiveness. Annual Review of Environmental Resources, 34, 305-324.

Boyd, S. (2013). Financing and Managing Energy Projects Through Revolving Loan Funds. Sustainability, 6(6), 345-352.

Emmenegger, P. (2011). Job security regulations in Western democracies: A fuzzy set analysis. European Journal of Political Research, 50(3), 336-364.

Fiss, P. (2011). Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research. Academy of Management Journal, 52(2), 393-420.

Fowler, K. M., & Rauch, E. M. (2006). Sustainable Building Rating Systems: Summary. Richland: Pacific Northwest National Laboratory.

Goertz, G., & Mahony, J. (2012). A Tale of Two Cultures. Princeton: Princeton University Press.

Mendel, J., & Korjani, M. (2013). Theoretical aspects of Fuzzy Set Qualitative Comparative Analysis. Information Sciences, 237(1), 137-161.

Policy and Society. (2013). Innovative methods for policy analysis: QCA and fuzzy sets. Policy and Society, 32(4), 279-356.

Ragin, C. (1987). The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies. Berkeley: University of California Press.

Ragin, C. (2008). Redesigning Social Inquiry: Fuzzy Sets and Beyond. Chicago: Chicago University Press.

Rihoux, B. (2013). Qualitative Comparative Analysis (QCA), Anno 2013: Reframing The Comparative Method’s Seminal Statements. Swiss Political Science Review, 19(2), 233-245.

Rihoux, B., & Ragin, C. (2009). Configurational Comparative Analysis. London: Sage.

Schneider, C., & Wagemann, C. (2012). Set-Theoretic Methods for the Social Sciences. Cambridge: Cambridge University Press.

Thøgersen, J. (2010). Country Differences in Sustainable Consumption. Journal of Macromarketing, 30(2), 171-185.

Van der Heijden, J. (forthcoming, 2014). Governance for Urban Sustainability and Resilience: Responding to Climate Change and the Relevance of the Built Environment. Cheltenham: Edward Elgar.

Verweij, S., Klijn, E.-H., Edelenbos, J., & Van Buuren, A. (2013). What makes governance networks work? A fuzzy set qualitative analysis of 14 dutch spatial planning projects. Public Administration, 81(4), 1035-1055.

Vis, B. (2012). The Comparative Advantages of fsQCA and Regression Analysis for Moderately Large-N analyses. Sociological Methods and Research, 41(1), 168-198.

Warren, J., Wistow, J., & Bambra, C. (2013). Applying Qualitative Comparative Analysis (QCA) to evaluate a public health policy initiative in the North East of England. Policy and Society, 32(4), 289-301.