Demand-Side Management Policy: Failure and Transferability

Peter Warren, University College London, London, UK

Abstract

The policy side of demand-side management (DSM) has received less attention in the academic literature than studies that focus on technological trials, utility programmes or modelling the future potential of DSM. The research contributes to filling this knowledge gap by undertaking a global systematic review of high-quality DSM policy evaluations conducted since the energy crises of the 1970s. The research aims to determine the mechanisms behind the implementation, success, failure and transferability of DSM policies introduced in all sectors (residential, commercial and industrial, but excluding transport and agriculture). However, the results presented in this paper focus on the latter two aims – policy failure and transferability.

The systematic review synthesises 119 high-quality documents covering 690 evaluations, which were obtained from 35 academic, industrial and governmental databases that publish research on DSM. The results show that 35 countries and 61 states (including provinces and regions) across six continents have implemented DSM policies with at least one high-quality evaluation. Systematic reviews involve rigorous assessment of the methodological quality of documents included in the sample, and the research developed an assessment scale that can be used to apply the method to the energy policy field.

24 key failure factors were identified inductively from the systematic review and a technique was developed to analyse DSM policy success, which triangulates factor frequency and factor weighting analyses. The most important failure factors across policies and countries were a lack of monitoring during the implementation and evaluation of policies, and technical issues such as programme management issues and technological performance problems. The paper breaks down the key failure factors by DSM policy and by country/state. Successful DSM policies (including policy packages) are determined by country/state and a framework is presented to analyse the transferability of successful policies between countries/states. Four crucial contextual factors were analysed: electricity market structure, climate, energy demands, and electricity system structure. Three groups of countries/states were identified where there is a higher probability of successful transferability within the groups in comparison with those outside of the groups.

Overview

Balancing Competing Energy Policy Goals

Governments around the world are increasingly trying to balance competing energy policy objectives for energy security, affordability and carbon emissions reduction. In some regions, such as Africa and east-Asia, energy access and local pollution reduction are still prominent goals. This challenge has been fuelled by climate change issues, resource depletion and consequent price increases, the growth of electronics in society, geopolitical issues, and national energy demand growth, particularly in emerging economies such as China and India.

To balance competing energy policy goals, governments are looking to the enhanced deployment of low carbon power plants (such as renewables), increasing interconnections between countries, developing and building the required energy infrastructure for a smart(er) grid, funding research into energy storage technologies (such as compressed air energy storage), and exploring demand-side options. All of the options are important in meeting this challenge and the paper focuses on the complementary role of the latter solution, demand-side management (DSM).

Demand-Side Management (DSM)

DSM refers to "technologies, actions and programmes on the demand-side of energy meters that seek to manage or decrease energy consumption, in order to reduce total energy system expenditures or contribute to the achievement of policy objectives such as emissions reduction or balancing supply and demand" (Warren, 2014a). DSM includes energy efficiency (such as insulation or efficient lighting), demand response (such as time-of-use pricing or incentive payments for peak load reduction), and on-site back-up generation and storage (such as micro-generation or hot water storage tanks). DSM policies can be classified into: 'Policy Type', 'Policy Category', and 'Specific Policy', with each level becoming more detailed (Warren, 2014c). Only the first two levels are given in full below, as the research analysis takes place at the 'Policy Category' level. However, some examples at the 'Specific Policy' level are given in brackets:

Market-based:
- Incentive payment-based demand response (e.g. interruptible/curtailment programmes)
- Price-based demand response (e.g. time-of-use pricing, critical peak pricing tariffs)
- Market transformations (e.g. removal of market barriers, market stimulation programmes)

Regulatory:
- Infrastructure rollouts (e.g. smart meter rollouts, energy display monitor rollouts)
- Utility obligations (e.g. energy efficiency resource standards, white certificate schemes)
- Labelling (e.g. appliance energy efficiency labelling, building labelling)
- Performance standards (e.g. equipment energy efficiency standards, building codes)

Fiscal:
- Loans and subsidies (e.g. tax incentives, grants, low-interest loans)
- Utility business models (e.g. decoupling policies, integrated resource planning)
- Research and development programmes (e.g. funding for deployment programmes)

Information-based:
- Information campaigns (e.g. energy audits, training programmes, education campaigns)

Voluntary:
- Voluntary programmes (e.g. industrial voluntary agreements, large commercial agreements)

The research analyses the twelve policy categories listed above to explore the implementation, success, failure and transferability of DSM policies. A further nine policy packages (of the above policies) are also included in the analysis in the second part of the paper. The concept of DSM in policy can be traced back to the USA's *National Energy Conservation Policy Act* and *Public Utility Regulatory Policy Act* (PURPA), which were introduced as part of the *National Energy Act 1978* (McNerney, 1998, pp. 27). This was the first instance of DSM being legislated nationally as a solution to the energy policy challenges of the 1970s. Nevertheless, the *notion* of DSM has been around for a long time, traditionally referring to a utility's general load management (Gellings and Chamberlin, 1993) or through the use of hot water tanks and off-peak storage heaters in houses (Barrett, 2006). The latter was particularly the case in New Zealand and Europe in the 1960s and 1970s (Gellings, 1985). However, PURPA was the first instance of DSM in government policy.

Research Overview

Since the 1970s there has been a vast literature on DSM, though the focus has primarily been on technological trials, utility programmes, and modelling studies on the technical potential of DSM.

DSM policy evaluation has received much less attention and is often limited to assessments of expected impacts rather than post-policy evaluations (Warren, 2014c). This paper presents some of the results from a three-year study that aimed to collate and synthesise all of the high-quality DSM (post-) policy evaluations that have been conducted around the world. The four aims of the research are to examine DSM policy implementation, to determine the key factors behind policy success and policy failure, and to explore the transferability of successful DSM policies between countries. The paper focuses on the latter two aims – policy failure and transferability, as the first two aims have been discussed elsewhere in Warren (2014c).

Methodology

Systematic Reviews: Theory

The research fits into the pragmatism research philosophy, which focuses on the method and using methods that are well suited to meeting the research aims (Saunders *et al.*, 2007). To answer higher level questions in a field where DSM research has often been narrow in focus (with the exception of modelling studies of the future potential of DSM), a systematic review is employed. The method has had limited application outside of the medical sciences, where it is commonly used to obtain a holistic understanding of the current evidence base. The Cochrane Collaboration provides a database of >5,000 systematic reviews in the medical sciences. However, the Campbell Collaboration, established in 2000, is beginning to apply the method to other policy areas, such as education, crime and justice, and social welfare. Nevertheless, it has had limited application in the energy policy field and there have been calls for this to be undertaken (Sorrell, 2007; Watson *et al.*, 2011). This research contributes to filling this methodological gap.

Systematic reviews involve collating all of the studies that have been done on a particular intervention and aggregating or synthesising the data in order to better understand the outcomes of the intervention (Warren, 2014b). Documents that may be included in a systematic review are: published and unpublished material, academic and 'gray' literature (such as policy documents, consultancy reports, and industrial reports), peer reviewed and non-peer reviewed documents, and English and non-English publications. However, a key part of the method, unlike traditional reviews (such as literature reviews or quick scoping reviews), is the detailing of a search strategy, inclusion and exclusion criteria, and the critical appraisal of the methodological quality of the documents reviewed (Petticrew and Roberts, 2006). As such, systematic reviews are a rigorous *method* that aims to gather data and evidence that can be analysed, unlike traditional literature reviews.

There are different types of systematic review, which fall into the following three board categories: integrative (such as meta-analysis), interpretive (such as qualitative synthesis), and mixed methods (such as realist synthesis). Integrative systematic reviews primarily involve statistical aggregation (though they can have qualitative elements), whereas interpretive systematic reviews primarily involve qualitative analysis (though they can have quantitative elements). The research presented in this paper utilised a mixed methods systematic review, specifically *realist synthesis*, as it is the most appropriate method for understanding how and why interventions work (or do not) by exploring their underlying mechanisms (Pawson, 2002a, Pawson, 2002b). Other types of systematic review, such as meta-analyses can be useful for assessing the quantitative performance of policies, but they cannot say how and why policies worked or failed, and are dependent on the availability of reliable quantitative data, which (unlike the medical sciences) are often difficult to find in the energy policy field in the numbers required for statistical aggregation (as this research discovered).

Systematic Reviews: Practice

All types of systematic review involve four filtering stages: firstly the number of initial hits from inputting a relevant search term into a database is recorded; secondly the titles of the hits are scanned for relevant documents; thirdly the abstracts are read to further reduce the sample size; fourthly the full document is read and subjected to a study quality assessment scale. The documents that reach the quantitative score threshold for passing the assessment scale are included in the sample.

The systematic review included 119 high-quality DSM policy evaluation documents from 35 databases (eleven academic, ten industrial, thirteen policy databases, and one interviews database covering quantitative interviews with seventeen DSM experts). The initial number of total hits aggregated across all of the databases was 15,894 documents. Due to the nature of systematic reviews to capture all relevant evidence of a high quality, the total number of documents is not pre-defined and instead the sample size is determined by the amount of existing evidence. It is important to note that the 119 documents did not equal 119 policy evaluations, as the majority of the documents included comparative analyses of different countries and policies. Thus, the final number of evaluations (which varied in their depth of analysis) totalled 690. All 119 documents passed a proposed scale for assessing study quality in energy policy evaluations (the *Warren Scale* – see Warren, 2014b). The scale is shown in table one below.

Table 1. *Warren Scale*: assessing study quality in energy policy evaluations (Warren, 2014b)

4 points (Implementation)	Who implemented the policy?
	How was the policy designed?
	How was the policy implemented?
	Why was the policy implemented?
4 points (Evaluation)	Who evaluated the policy?
	How was the policy evaluated?
	What were the impacts of the policy?
	How successful was the policy?
2 points (Peer Review)	Was the evaluation peer reviewed by a relevant expert?
	Was the evaluation peer reviewed by 2+ relevant experts?
2 points (Acknowledgements)	Are there statements of conflicts of interest and/or copyright?
	Are there statements of acknowledgement?
1 point (Reliability)	Is the author/institution reliable?
1 point (Context)	Are the totals given when percentages are used?

Documents must reach half of the total number of points to be included in the final sample (7 out of 14 points). The scale was developed from investigating the assessment scales used in other disciplines. However, most focus on whether or not the study was randomised, double blind, and described the withdrawal rates (Jadad, 1998), aspects that are more suited to medical trials than DSM policy evaluations, which rarely conduct randomised control trials due to practical and ethical reasons. The *Warren Scale* is developed to be better adapted to the types of evaluations found in the energy policy field where the quantitative aspects primarily revolve around policy impacts (based on specific performance criteria, such as energy savings, costs to government, carbon emissions reduction, etc.) and have large qualitative aspects (such as discussions of performance).

Approach for Data Analysis

The research does not use a specific definition for 'policy success', as this varied across the evaluations included in the systematic review. Instead, it encompasses all of the definitions in the sample (such as meeting performance criteria or determining whether or not original policy objectives were met), and uses a proposed quantitative technique to analyse policy success in order to establish whether or not a DSM policy has been successful in a particular country or state. States (including provinces and regions) are included in the analysis in addition to countries in order to expand the evidence base. A large number of DSM policy evaluations have been conducted at state-level rather than at national-level, particularly in the USA.

To investigate the first aim of the paper, DSM policy failure, the proposed approach for data analysis focuses on conducting frequency analysis of the key failure factors across evaluations, undertaking weighting analysis of the key failure factors within evaluations (for every document), and then triangulating the two analyses to get an accurate picture of the most important failure factors generally, and broken down by DSM policy and country/state. Firstly, in order to calculate the frequency of failure factors by DSM policy, a threshold is used to categorise factors into those that have a high frequency of discussion across evaluations for a given policy, and those that have a low frequency of discussion. The threshold was determined inductively based on how frequent factors tended to be discussed in the documents (~5 evaluations was considered average):

Factor Frequency Threshold:

1) **High Frequency:** ≥5 evaluations
2) **Low Frequency:** <5 evaluations

Those policies where a given factor was discussed in ≥5 evaluations were considered important. However, by itself, frequency can only give an indication of a factor's importance *across* evaluations, which does not necessarily indicate its importance *within* evaluations. Thus, the weighting of factors within evaluations is also needed in the determination of policy success and failure. In order to calculate the weightings of failure factors by DSM policy, a 1.0-3.0 weightings scale was used for each evaluation (that had enough depth of analysis) within each document in the analysis. The scale is based on the qualitative emphasis that evaluators give to various failure factors. To reduce the subjectivity of converting qualitative statements on the emphasis of factors into quantitative data, specific words of emphasis were examined:

Factor Weighting Scale:

1) ***Score weighting 2.5-3.0 (Crucial):*** the following words are used in direct relation to the factor to strongly emphasis its importance: 'critical', 'crucial', 'very important', 'necessary', 'primary reason(s)', 'key', 'vital', 'central', 'essential', 'fundamental', 'imperative', 'decisive', 'significant', or equivalent
2) ***Score weighting 1.5-2.4 (Some importance):*** the factor is included at the start of a list and is frequently discussed though it is not strongly emphasised using any of the words for score weighting 3, or it is referred to using phrases such as: 'quite important', 'had some influence', 'played a role', or equivalent
3) ***Score weighting 1.0-1.4 (Small impact but not unimportant):*** the factor is included towards the middle or end of a list of other factors without emphasis or discussion, or it is indirectly inferred as a factor
4) ***No weighting (Unimportant):*** no weighting is given to the success factor as it is considered unimportant to warrant discussion

Initially a factor is given a score of 1, 2 or 3 from the scale (where 3 is the highest weighting). For the analysis of failure factors by policy, averages of the weightings are calculated across countries and states for each policy. For the analysis of failure factors by country or state, averages of the weighting are calculated across policies for each country/state. This produces a figure to one decimal place, thus converting the weighting into the 1.0-3.0 scale above. Weightings give an indication of the importance of factors in a given context (i.e. a specific policy implemented in a particular country/state) but cannot indicate whether or not the pattern is found in other contexts. As such, frequency is needed to overcome this drawback, and thus, the weakness of each technique is the strength of the other, and this justifies the need to triangulate the two methods.

The equations presented below propose how to triangulate frequency and weighting analyses:

Frequency-Weighting Triangulation Equation for Policy Failure:

1) Frequency-Weighting triangulation (FW_{trg}) = Policy Failure weighting (PF_w) x (Policy Failure factor frequency in sample (PF_{ffs}) x Policy Failure factor weighting (PF_{fw})) / 10

2) Frequency-Weighting Triangulation Score (FWS_{trg}) = (Frequency-Weighting triangulation (FW_{trg}) / Theoretical Maximum Triangulation Score (FWS_{trgmax})) x 100

In notation form:

1) $FW_{trg} = PF_w \text{ x } (PF_{ffs} \text{ x } PF_{fw}) / 10$

2) $FWS_{trg} = (FW_{trg} / FWS_{trgmax}) \text{ x } 100$

The first part of the equation calculates the triangulation values for each factor for each DSM policy. Dividing by ten at the end of the calculation is required in order to produce figures for a more comparable and manageable scale for categorising factors. The second part of the equation then converts these values into a proposed scale to categorise the values into crucial factors, important factors, and unimportant factors for each DSM policy and then for each country/state, as shown below. Multiplying by 100 at the end of the calculation is required to convert the figures into percentages so it is expressed as a percentage of the theoretically maximum triangulation score that could be achieved (~5% of the theoretical maximum was inductively determined to be average):

Factor Triangulation Scale:

1) ≥10.0% of theoretical maximum = Crucial factor
2) 5.0-9.9% of theoretical maximum = Important factor
3) <5.0% of theoretical maximum = Unimportant factor

The theoretical maximum is calculated by taking the overall weighting of a given DSM policy (across countries/states) in the systematic review sample, multiplying it against the multiplication of the frequency of discussion of that policy in the sample against the maximum possible success weighting of the policy. The value is then divided by ten for the same reasons as the first part of the triangulation equation (in order to produce figures for a more comparable and manageable scale for categorising factors). The equation is more easily understood in notation form:

Theoretical Maximum Equation:

1) Theoretical Maximum Triangulation Score (FWS_{trgmax}) = Policy Success weighting (PS_w) x (Policy frequency in sample (P_{fs}) x Theoretical Maximum Policy Success Weighting (PS_{wmax})) / 10

In notation form:

1) $FWS_{trgmax} = PS_w \text{ x } (P_{fs} \text{ x } PS_{wmax}) / 10$

The 'Policy Success weighting (PS_w)' is calculated using a scale of 1-5 (not 1-3 as with the factor weighting scale) and determines how successful a policy is in a given country or state. Averages are then calculated across countries/states for the analysis of policies. Note that PS_w becomes CS_w ('Country Success weighting') when the focus switches to failure factors by country/state rather than failure factors by policy. However, all equations are the same for both types of analysis. This is discussed further in the results section. A final important point to note is that one of the research aims that is not discussed in this paper is the analysis of *success* factors by DSM policy and by country/state. The analysis is conducted using exactly the same method as for *failure* factors by policy and by country/state. However, both are needed in the calculation of PS_w or CS_w.

To investigate the second aim of the paper, DSM policy transferability, data is obtained for each country/state in the sample on the most important contextual factors that influence it: electricity market structure, climate, energy demands, and electricity system structure. Their importance was judged based on a literature review conducted prior to the systematic review. Data were obtained from the International Energy Agency (IEA) and the Energy Information Administration (EIA). At the time of the analysis, 2011 data was the most recent data available. More specifically, the data sources were: the IEA's *Country Statistics* and *Energy Policies of IEA Countries Reviews*, and the EIA's *Country Statistics*. A transferability framework was developed to store the information, and once successful DSM policies had been identified by country/state, this was added to the framework. Countries or states that matched together by context (similar market structures, climates, energy demands, and system structures) were placed together in a 'transferability group'. It is then assumed that if any of those countries or states have had success with any of the twelve individual policies or nine policy packages analysed, there is potential for it to be transferred to the other countries/states in the group. However, it is noted that there are other important contextual factors that could not be included in the transferability analysis, such as regulatory and political structure, consumer (historical) familiarity with DSM, the degree of environmentalism, and the history of DSM policy implementation. This was primarily due to difficulties in obtaining data for all 35 countries and 61 states in the sample. However, the literature review and the systematic review have highlighted the importance of market structure as the most crucial factor influencing transferability, and this was a detailed part of the analysis. Nevertheless, the transferability framework has been designed to be a user-friendly tool, which can be easily updated to include new data that becomes available for both factors already included and new contextual factors.

The electricity market structure factor is broken down into five sub-factors: generation (state-owned, partially-privatised, or fully-competitive – IEA and EIA 2011 data), transmission and distribution (state-owned, partially-privatised, or fully-competitive – IEA and EIA 2011 data), the presence of DSM in balancing or reserve markets (yes, no, or forthcoming – data from government and utility websites), utility structure (vertical or horizontal – IEA and EIA 2011 data), and the presence of submitting DSM plans to regulators (yes or no – IEA DSM programme 2012 data). The climate factor is broken down into two sub-factors: the historical summer average maximum temperature (°C) and the historical winter average minimum temperature (°C). 1990-2009 averaged data was obtained from the World Bank. Projected summer and winter temperatures can be added to the framework, though this was not done for the current analysis. The energy demands factor is broken down into two sub-factors: electricity consumption (GWh) and heat consumption (TJ) (IEA 2011 data). The electricity system structure factor is broken down into four sub-factors: fossil fuels (%), nuclear (%), hydro (%), and other renewables (%). The figures represent the percentage of the electricity system. Other renewables refers primarily to wind power and biomass burning. Only national data were available for the energy demands and electricity system structure factors, which may reduce the output for the number of states where transferability could be feasible. The transferability analysis involved 112 iterations of the contextual factors. Analysis took place at the sub-factor level where each combination of the sub-factors against other sub-factors was undertaken.

Results

Key Failure Factors

The results for the analysis of DSM policy failure were split into three parts: the overall key failure factors across policies and countries/states, the key factors by policy, and the key factors by country/state. This is an area of current analysis and the final results will be available for presentation in the conference. As such, only the first two parts are presented here. All of the results follow the application of the triangulation equations to the frequency and weighting analyses.

The results of the triangulation analysis show that across policies and countries/states, lack of monitoring and technical issues are the two the most important failure factors. By frequency, they were discussed in 41% and 37% of the documents respectively (49/119 and 44/119 documents respectively). By weighting, the factors score 2.0 and 1.8 respectively and have the highest weightings of any factor in the analysis. Lack of monitoring refers to inadequate (comprehensive) evaluation during and after the policy period to monitor policy performance. Technical issues refer to programme management issues, technological performance problems, and a lack of required technical skills. At the other end of the scale, incentivising utilities and stakeholders and utility opposition appear to be less important than the other factors in the analysis. However, this does not suggest that they are unimportant for specific policies or countries/states as the analysis below shows. It is important to note that all of the factors in the analysis should be managed in the design and implementation of DSM policies, as these have been inductively determined from the systematic review as having some importance rather than being pre-defined. Nevertheless, the analysis aims to show how the factors compare against each other. Figures one and two summarise the results by frequency and weighting. Factors are categorised by colour into groups: policy issues, consumer issues, political issues, stakeholder engagement, and infrastructure issues, and the same format is used for both graphs.

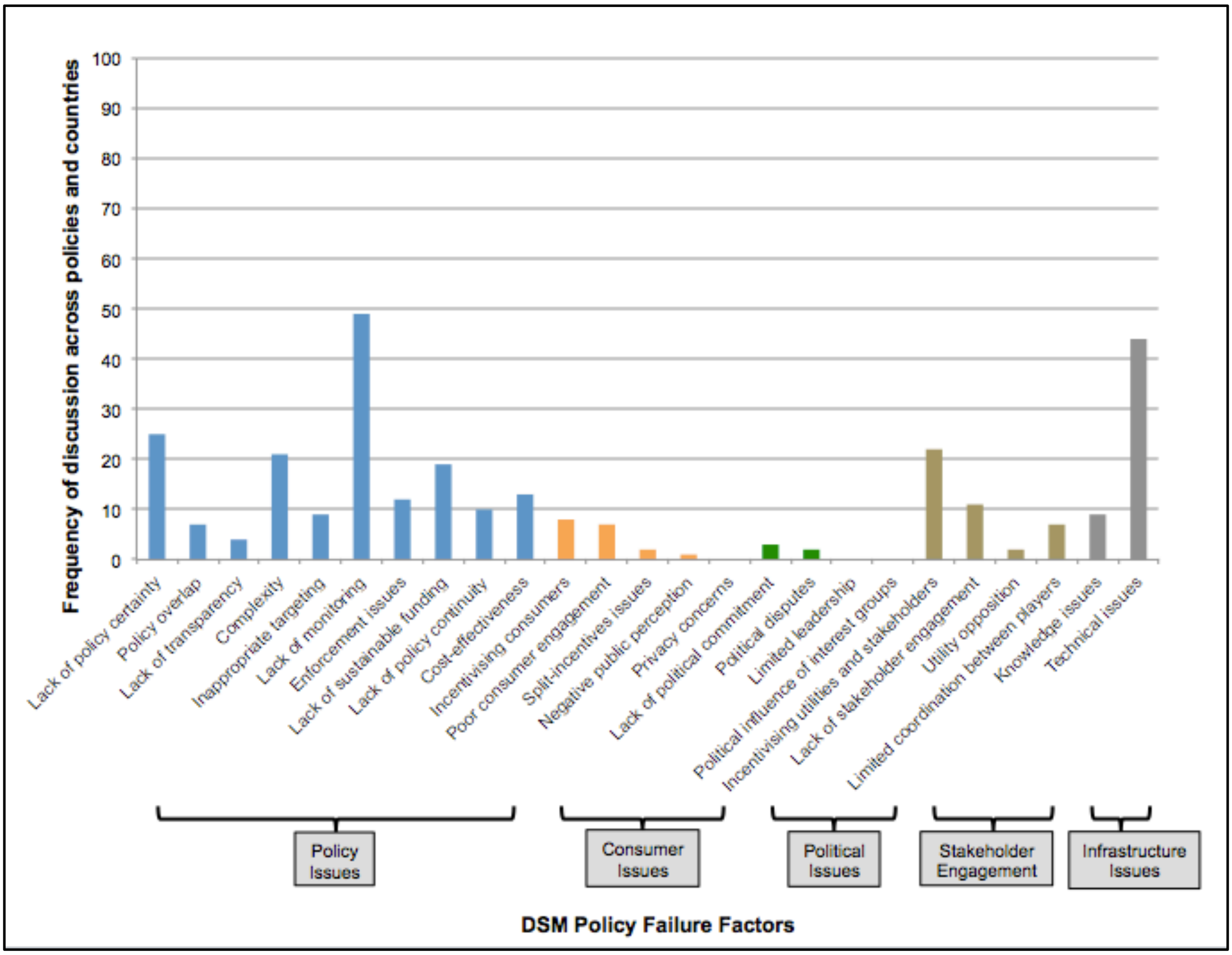

Figure 1: Frequency of DSM policy failure factors across policies and countries/states

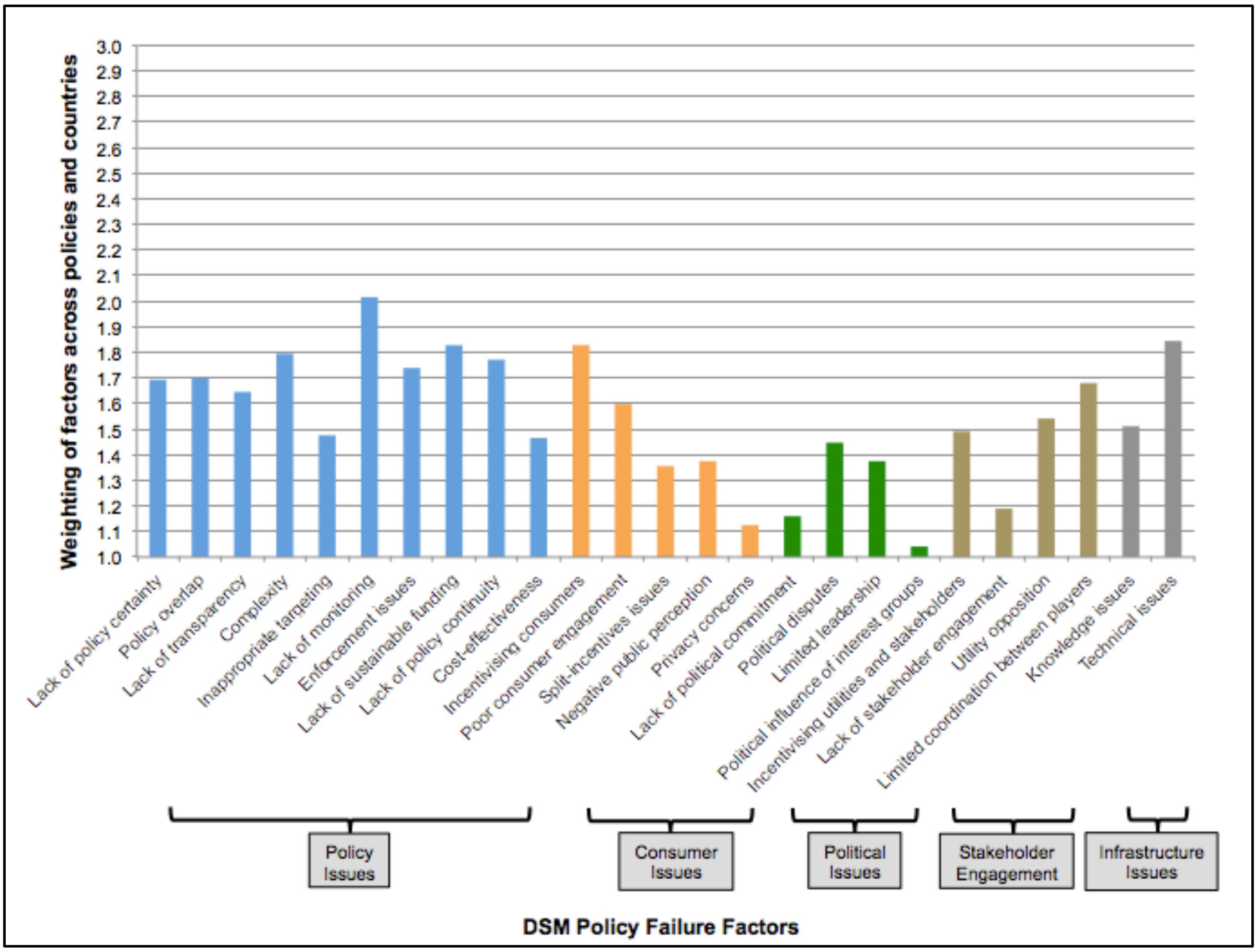

Figure 2: Weighting of DSM policy failure factors across policies and countries/states

When the results are looked at by DSM policy, the following factors are crucial (underlined and in italics) and important (just in italics):

- **Incentive payment-based demand response:** *lack of policy certainty, technical issues*
- **Price-based demand response:** *lack of monitoring*, *incentivising consumers*, *incentivising utilities and stakeholders*, *technical issues*
- **Market transformations:** *lack of policy certainty*, *lack of monitoring*, *technical issues*
- **Infrastructure rollouts:** <u>*lack of policy certainty*</u>, <u>*policy overlap*</u>, <u>*lack of transparency*</u>, <u>*political disputes*</u>, <u>*limited leadership*</u>, <u>*incentivising utilities and stakeholders*</u>, <u>*limited coordination between players*</u>, <u>*technical issues*</u>, *lack of policy continuity, negative public perception, privacy concerns, knowledge issues*
- **Utility obligations:** (no factors pass triangulation threshold)
- **Labelling:** *technical issues*
- **Performance standards:** (no factors pass triangulation threshold)
- **Loans and subsidies:** (no factors pass triangulation threshold)
- **Utility business models:** (no factors pass triangulation threshold)
- **R&D programme:** <u>*technical issues*</u>, *incentivising consumers*
- **Information campaign:** (no factors pass triangulation threshold)
- **Voluntary programmes:** (no factors pass triangulation threshold)

It is clear that the findings by policy show similarities to figures one and two with technical issues dominating. Lack of monitoring also scores highly and is considered important alongside lack of policy certainty and incentivising consumers. Lack of policy certainty refers to the presence of unclear timeframes, a lack of clarity on certain policy details, and limited legislation to guarantee the implementation of the policy. Incentivising consumers refers to a lack of adequate and appropriate consumer incentives that are clearly communicated. Where no factors are listed for certain policies, no factors passed the triangulation threshold (of ≥10% of the theoretical maximum). However, if the frequency and weighting analyses are looked at separately, all of the policies have a list of crucial and important factors.

Successful DSM Policies

The research proposes an equation for calculating DSM policy success based on the underlying policy mechanisms. This is determined by analysing the results for the frequency and weighting of failure *and* success factors across all policies and countries/states. In this paper, only the analysis of failure factors is presented, as the analysis of success factors has been presented elsewhere in Warren (2014c). This is an area of current analysis and the final results will be available for presentation in the conference, and as such, detailed discussion of the equation is not given here. However, in short, the equation produces figures to one decimal place between 0.0-5.0 (though theoretically, scores higher than 5.0 could be achieved). DSM policies scoring >3.5 overall (across countries and states in the sample) are considered successful and those scoring ≤3.5 are considered unsuccessful or have performed below average.

When the equation is applied to the systematic review data, the following results are produced, as summarised in table two. In this stage of the analysis, policy packages are included due to data availability. The table shows a list of policies that have been successfully implemented and evaluated in specific countries and states.

Table two. Successful DSM policy implementation and evaluation by country/state

Country or State	Successful Policies
USA (California)	IPBDR, PBDR, MT, IR, PS, L&S, UBM, R&D, IC, UBM/MT
China	IPBDR, PS, VP, PS/LB, IC/L&S, PBDR, IC, LB, L&S, R&D
UK	IPBDR, PBDR, IR, UO, PS, L&S, UBM, R&D, IC
USA	IPBDR, PBDR, UO, PS, L&S, UBM, R&D, UBM/MT, IC/L&S
Denmark	UO, LB, PS, L&S, R&D, IC, VP
Thailand	MT, LB, L&S, IC
USA (New York)	IPBDR, L&S, UBM, UBM/MT
USA (Vermont)	PBDR, UO, PS, UBM
USA (state-level)	UO, PS, UBM
Germany	L&S, IC, IC/L&S
USA (PJM region)	PBDR, IPBDR/PBDR
France	PBDR, UO
USA (Pacific Northwest region)	UBM/MT, PS/IC
USA (Massachusetts)	UBM/MT, IC/L&S
European Union (EU)	UO, PS
Australia	IR, PS
USA (NYISO region)	IPBDR/PBDR
China (Hebei)	UBM
China (Fujian)	UBM

Belgium (Flanders)	UO
Italy	UO
Japan	UO
Brazil	UO
Estonia	L&S
USA (ISO-NE region)	IPBDR/PBDR
USA (Illinois)	IC/L&S
USA (Florida)	IPBDR
China (Jiangsu)	IPBDR/PBDR
China (Beijing)	IPBDR/PBDR
India (Orissa)	L&S
Australia (New South Wales)	UO
Australia (Australian Capital Territory)	UO
Australia (South Australia)	UO
Australia (Victoria)	UO
Philippines	PS/LB/IC
USA (Ohio)	UBM
Spain	IPBDR
USA (Wisconsin)	IC/L&S
Canada	UO
South Korea	IC
Sweden	MT

Table acronyms:

Individual DSM Policies:

- IPBDR: Incentive payment-based demand response
- PBDR: Price-based demand response
- MT: Market transformations
- IR: Infrastructure rollouts
- UO: Utility obligations
- LB: Labelling
- PS: Performance standards
- L&S: Loans and subsidies
- UBM: Utility business models
- R&D: Research and development programmes
- IC: Information campaigns
- VP: Voluntary programmes

DSM Policy Packages:

- IPBDR/PBDR: IPBDR and PBDR policy mix
- UBM/MT: UBM and MT policy mix
- IC/L&S/MT: IC, L&S and MT policy mix
- PS/LB/UO/L&S: PS, LB, UO and L&S policy mix
- PS/LB/IC: PS, LB and IC policy mix
- PS/LB: PS and LB policy mix
- IC/L&S: IC and L&S policy mix
- PS/IC: PS and IC policy mix

- VP/L&S: VP and L&S policy mix

As table two highlights, California (USA), China, the UK, and the USA have had the most success and diversity with DSM policies, having successfully implemented and evaluated 9-10 individual policies or policy packages. Denmark also performs well with seven successfully implemented and evaluated DSM policies. Furthermore, it is clear that utility obligations (UO) have the widest spatial distribution with sixteen countries/states having successfully implemented and evaluated the policy. Performance standards (PS), loans and subsidies (L&S), and utility business models (UBM) also have wide spatial distributions with nine countries having successfully implemented and evaluated the policies. The top four policies are frequently discussed across evaluations in the systematic review and are generally highly weighted within evaluations. Overall, individual DSM policies appear to perform better than DSM policy packages. One of the primary purposes of identifying successful DSM policies by country/state is for the transferability analysis.

Transferability Potential

The second part of the analysis aimed to explore the potential transferability of successful DSM policies between countries/states. Data were obtained for four of the most important contextual factors affecting transferability: electricity market structure, climate, energy demands, and electricity system structure, and inputted into a transferability framework. The successful DSM policies by country/state identified in the previous section were also inputted into the framework. 112 different iterations were undertaken to group countries/states with the same electricity market structures, climates, energy demands, and electricity system structures. The framework makes the assumption that if any of the countries/states within a group have successfully implemented a particular DSM policy, then there is a higher probability that it could have similar success in the other countries/states in the group compared with those not in the group.

The results produce three groups of countries/states with similar enough contexts for potential transferability. In the first group, much of what has been done at either a state-level or national-level in the USA could be transferred within the country. Japan is also included here and there is the potential for a large number of policies (IPBDR, PBDR, UO, PS, L&S, UBM, R&D, UBM/MT, and IC/L&S) to be transferred from the USA to Japan. The second group has the largest transferability potential (in terms of numbers of countries/states in the group) and includes Austria, Belgium, Finland, Germany, New Zealand, Sweden, and the UK. Sweden's success with market transformations (MT) could be transferable within the group, as could Germany's successful experiences with information campaigns and the policy package of loans and subsidies with information campaigns. However, the UK provides the largest potential for transferring policies to the other countries in the group with nine successful policies (IPBDR, PBDR, IR, UO, PS, L&S, UBM, R&D, and IC). The third group involves just two countries – Australia and the Philippines. Here there is potential for Australia's successful experiences with infrastructure rollouts (IR) and performance standards (PS) to be transferred to the Philippines, and for the successful implementation of the PS/LB/IC policy package (performance standards with labelling and information campaigns) in the Philippines to be transferred to Australia.

What is equally interesting is where it is unlikely that successful DSM policies could be transferred. The groupings above show strong potential for transferability, however, outside of this, the potential is much weaker. For example, if only the two most important contextual factors in the research are examined (electricity market structure and climate) and the other two factors are excluded from the analysis (energy demands and electricity system structure), the number of countries where transferability could be successful increases. Here, eight groups of similar countries/states are produced. However, as the number of contextual factors included has reduced, the transferability potential should be considered weaker than those in the original three groups. Outside of these eight groups (or the original three groups) the potential for transferability is limited.

Conclusion

Demand-side management (DSM) policy is an under-researched area with the focus having been on technological trials, utility programmes, and studies modelling the technical potential of DSM. The primary aims of the presented research are to examine DSM policy implementation, to determine the key factors behind policy success and failure, and to explore the transferability of successful DSM policies between countries/states. This paper focussed on the last two aims – the failure and transferability of DSM policies. A systematic review is employed to synthesise high-quality DSM policy evaluations conducted around the world, and a method for examining the factors behind policy failure is proposed, which uses the analysis of both frequency and weighting and then triangulating the results.

The research synthesised 119 high-quality documents (covering 690 evaluations) from 35 academic, industrial and governmental databases. 35 countries and 61 states (including provinces and regions) across six continents were included, and 12 individual DSM policies and 9 policy packages were examined. 24 failure factors were identified, and the results showed that overall, across policies and countries/states, lack of monitoring and technical issues were the two most important factors contributing to the failure of DSM policies (or causing policies to perform less well than originally anticipated). When the key failure factors by DSM policy are examined, six of the twelve policies in the analysis show crucial and important factors that if unmanaged, are likely to cause a policy to fail. A high proportion of these policies have technical issues, lack of monitoring, lack of policy certainty, and incentivising consumers as key factors. The key failure factors by country/state are currently being examined and the full results will be available for presentation at the conference.

The research identified successful DSM policies by country/state. The countries/states with the highest number of successful DSM policies are California (USA), China, the UK, and the USA with 9-10 DSM policies (including policy packages) having been successfully implemented and evaluated. Utility obligations have the widest spatial distribution with sixteen countries/states having successfully implemented the policy. Performance standards, loans and subsidies, and utility business models also have wide spatial distributions with nine countries having successfully implemented them. The identification of successful DSM policies allowed an analysis of the potential transferability of successful policies between countries/states. A quantitative transferability framework was produced to match up countries/states with similar contexts. The main contextual factors examined were electricity market structure, climate, energy demands, and electricity system structure. The findings suggest that there are three main groups of countries/states with similar enough contexts where the probability of transferability within each group is higher in comparison to those countries/states not in the groups. Group one consists of the USA (national-level), USA (state-level), and Japan, the second group consists of Austria, Belgium, Finland, Germany, New Zealand, Sweden, and the UK, and third group consists of Australia and the Philippines. Electricity market structure appears to be the most important factor, and the analysis has also shown where transferability may not be possible.

References

1. Barrett, M. (2006) *A renewable electricity system for the UK*, A response to the 2006 Energy Review, University College London (UCL), London, UK
2. Campbell Collaboration: www.cochrane.org
3. Cochrane Collaboration: www.campbellcollaboration.org
4. Energy Information Administration (EIA) (2011) *Country Reports*, can be downloaded from: http://www.eia.gov/countries/
5. Gellings, C.W. (1985) "The concept of demand-side management for electric utilities", *Proceedings of the IEEE*, 73 (10)

6. Gellings, C.W. and Chamberlin, J.H. (1993) *Demand-Side Management: Concepts and Methods*, 2nd Edition, The Fairmont Press, Inc., USA
7. International Energy Agency (IEA) (2011) *Country Statistics*, can be downloaded from: http://www.iea.org/statistics/
8. International Energy Agency (IEA) (2011) *Energy Policies of IEA Countries Reviews*, can be downloaded from: http://www.iea.org/publications/countryreviews/
9. International Energy Agency (IEA) (2012) *Best practices in designing and implementing energy efficiency obligation schemes*, Research Report: Task XXII of the IEA DSM Programme, prepared by the Regulatory Assistance Project (RAP)
10. Jadad, A. (1998) *Randomised controlled trials: a users guide*, London: BMJ Books, UK
11. McNerney, R.A. (1998) Changing Structure of the Electric Power Industry: An Update, Chapter 4, pp. 27-28, DIANE Publishing
12. National Energy Act 1978, Washington DC: National Congress, USA
13. National Energy Conservation Policy Act 1978, Washington DC: National Congress, USA
14. Pawson, R. (2002a) "Evidence-based policy – in search of a method", *Evaluation*, 8, 2
15. Pawson, R. (2002b) "Evidence-based policy: the promise of 'Realist Synthesis'", Evaluation, 8, 3
16. Petticrew, M. and Roberts, H. (2006) *Systematic Reviews in the Social Sciences*, Blackwell Publishing, Oxford, UK
17. Public Utility Regulatory Policy Act 1978, Washington DC: National Congress, USA
18. Saunders, M., Lewis, P. and Thornhill, A. (2007) *Research Methods for Business Students*, 4th Edition, Prentice Hall
19. Sorrell, S. (2007) "Improving the evidence base for energy policy: the role of systematic reviews", *Energy Policy*, 35 (3), 1858-1871
20. Warren, P. (2014a) "A review of demand-side management policy in the UK", *Renewable and Sustainable Energy Reviews*, 29, 941-951
21. Warren, P. (2014b) "The use of systematic reviews to analyse demand-side management policy", *Energy Efficiency*, 7 (3), 417-427
22. Warren, P. (2014c) "Demand-Side Management Policy: Implementation and Success", *International Energy Policies and Programmes Evaluation Conference*, 9-11th September 2014, Berlin, Germany
23. Watson, J. *et al*. (2011) "What are the major barriers to increased use of modern energy services among the world's poorest people, and are interventions to overcome these effective?" *Systematic Review CEE 11-004*, Draft Paper, CEE, UK
24. World Bank (2014) *Climate Change Knowledge Portal*, 1990-2009 averaged data from the Climate Research Unit, University of East Anglia