Department for
Business, Energy
& Industrial Strategy

EVALUATION OF THE ENERGY SAVINGS OPPORTUNITY SCHEME

Impact evaluation scoping report

October 2017

This document is available in large print, audio and braille on request. Please email enquiries@beis.gov.uk with the version you require.

EVALUATION OF THE ENERGY SAVINGS OPPORTUNITY SCHEME

Impact evaluation scoping report

Any enquiries regarding this publication should be sent to us at enquiries@beis.gov.uk.

This publication is available for download at www.gov.uk/government/publications.

Contents

Executive Summary

Ipsos MORI and University College London (UCL) were commissioned by the Department of Energy and Climate Change (DECC)[1] in April 2015 to undertake a process evaluation of the Energy Savings Opportunity Scheme (ESOS) and to conduct a scoping (or feasibility) study into the design of a full impact evaluation. The findings of the scoping study are presented in this report[2].

Overview of ESOS

Article 8 (4-6) of the EU Energy Efficiency Directive (2012/27/EU) requires the UK to mandate that all large undertakings (and smaller organisations part of a group including a large undertaking) undertake audits (unless exempt from the requirement to undertake an audit as a result of having a qualifying energy management system in place such as ISO 50001) of the energy used by their buildings, industrial processes and transport to identify cost-effective energy saving measures by 5th December 2015 and at least every four years thereafter. ESOS was developed by DECC to implement the Directive in the UK. Obligated organisations were required to notify compliance with the scheme administrator[3], and implementation of the energy saving measures identified was voluntary. The objectives of ESOS[4] are to:

- provide large enterprises with enterprise-specific information about how they can make energy savings;
- stimulate the take-up of cost-effective energy efficiency measures;
- minimise the cost to businesses of complying with the regulations; and,
- maximise the synergies with existing policies.

[1] The Department for Business, Energy and Industrial Strategy (BEIS) as of July 2016

[2] A separate report presenting process findings, and exploring early indications of impact, is available here: https://www.gov.uk/government/publications/energy-savings-opportunity-scheme-esos-evaluation-of-the-scheme

[3] The Environment Agency is the scheme administrator for the whole of the UK. Responsibility for compliance and enforcement rests with the Environment Agency in England and the equivalent devolved agencies.

[4] Impact Assessment: Energy Savings Opportunity Scheme, 2014. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323116/ESOS_Impact_Assessment_FINAL.pdf

Rationale for intervention and expected outcomes

The UK Energy Efficiency Strategy (2012)[5] highlighted that 196 TWh of socially cost-effective savings could be achieved from energy efficiency investments, 69 TWh of which could be realised from the commercial and industrial sectors. While organisations may have an incentive to implement energy efficiency projects where the expected cost savings exceed the investment costs, the Strategy highlights asymmetries in information and cultural barriers that would prevent them being taking forward without intervention from the public sector. ESOS responds to these issues by obliging organisations to acquire information on the costs and benefits of the available investments and raising visibility of energy efficiency issues by requiring the recommendations to be signed-off at Board-level.

The key outcomes and impacts anticipated from ESOS include:

- greater investment in energy efficiency measures (including behavioural initiatives);
- reduced energy and fuel consumption and lower costs to business; and,
- environmental benefits resulting from reduced greenhouse gas emissions (with associated reductions in future abatement costs and improvements in air quality).

Based on conservative assumptions regarding the behavioural effects of ESOS, DECC's Impact Assessment estimated that the scheme would result in energy savings of around £2.2bn. Improvements in air quality were valued at £320m, with the value of future abatement costs associated with reduced greenhouse gas (GHG) emissions valued at a total of £280m[6].

Evaluation background, aims and approach

The overall aim of the scoping study presented in this report was to assess the most appropriate approach to conducting a robust impact and economic evaluation of ESOS. Such an evaluation would quantify the effects of ESOS on obligated organisations and

[5] UK Energy Efficiency Strategy: The Energy Efficiency Opportunity in the UK, 2012. Available at: https://www.gov.uk/government/publications/energy-efficiency-strategy-2013-update (Accessed July, 2015)

[6] Impact Assessment: Energy Savings Opportunity Scheme, Department for Energy and Climate Change and Department for Transport, 2014. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323116/ESOS_Impact_Assessment_FINAL.pdf (accessed July 2015).

understand how any energy efficiency benefits have been realised. The feasibility study focused on addressing the following central challenges:

- how the intended outcomes could be measured;
- how far it is possible to establish a suitable comparison group; and,
- the analytical methods that could be applied to establish causality.

The findings and recommendations made in this report are based on: analysis of scheme management data, a review of secondary datasets and data-linking opportunities, and an appraisal of the available statistical and non-statistical approaches to assessing impact. The analysis was guided by an overall evaluation framework developed at the outset of the feasibility study, including a detailed 'Theory of Change' describing how ESOS was expected to produce its intended results.

This evaluation took place against the ongoing work by the scheme compliance bodies to bring eligible organisations into compliance. This produced additional challenges for the scoping study both for ascertaining the size of the obligated population and in understanding how best to link these organisations to other datasets to measure outcomes.

Measurement of ESOS outcomes

An impact evaluation will require longitudinal observations of the outcomes of interest to measure the changes produced by ESOS. The feasibility study considered how far it would be possible to measure the key outcomes of interest using existing data sources and where collection of new data may be required.

Overall, the study has demonstrated the value, and feasibility, of matching into the Annual Business Survey (ABS) and capturing primary data for some key outcomes of interest (firm-level productivity and expenditure and behavioural outcomes), but has identified more considerable issues around measuring energy consumption. The findings indicated that:

- **Annual Business Survey:** Ninety-six percent of ESOS-obligated organisations were linked to the Annual Business Survey (ABS). This will provide longitudinal observations of capital investment, energy and fuel expenditure, GVA productivity, sales and profit[7] at

[7] The ABS is an annual survey of around 50,000 large businesses in the UK undertaken for the Office for National Statistics to collect information feeding into the calculation of key national accounting measures. Similar outcomes data is available for SMEs through the Structure Database.

an organisational level. This will enable tracking of many of the outcomes of interest though it will not provide direct measures of energy consumption (and data only for large firms).

- **Energy Consumption:** Approaches to tracking energy use tested during this scoping study have included data linking to the Non-Domestic National Energy Efficiency Database (ND-NEED) and matching ESOS sites to raw meter data. The effectiveness of these approaches was limited by the quality of the address information. Overall this testing suggests that whilst it is likely to be possible to match some energy consumption data to most ESOS obligated organisations, this will not be complete at a site level. There are also potential alternatives to source proxy energy data; one example is the ABS which records energy spend data for large organisations (however, use of a dataset focusing on large organisations would in turn limit the available evaluation options).

- **Primary research:** The baseline survey conducted as part of the process evaluation of ESOS collected other key outcome measures including energy efficiency measures installed, associated capital spend, number of employees involved in energy management, levels of priority placed on energy efficiency at Board-level, and, implementation of behavioural measures. Consent to re-contact was given by 71% of respondents from organisations who at the time of the baseline survey had notified compliance, which would allow for their outcomes to be tracked in follow-up surveys. There will also be value in gathering qualitative evidence to capture uptake of behavioural energy efficiency measures, perceptions of reputational impacts from ESOS compliance (or associated uptake of measures and strategies), and to explore future investment and energy consultancy plans to inform estimations on the potential effects of ESOS on suppliers of energy efficiency services and technologies.

It should be noted that administrative data will only be available with a lag. This will have timescale implications for a future quantitative study: it may not be possible to fully determine the effect of ESOS in 2017 until 2019/20.

Establishing a comparison group

A credible assessment of impact requires the identification of an appropriate comparison group to attribute changes in the key outcomes of interest to ESOS. A significant challenge arises from the universality of the ESOS obligations (there are no large organisations that are not exposed to the intervention from which a comparison group can be formed).

Overall, while the feasibility study identifies potential options for this comparison group, the selection between them requires a balancing of the generalisability of findings across the ESOS population (offered by an approach based on who had and had not yet notified

compliance) against selection bias (avoided through an approach based on non-obligated SMEs).

Given this constraint, two main possibilities have been identified:

- **SMEs:** SMEs that are not obligated under ESOS could serve as a comparison group for those SMEs that are (by virtue of being a member of an enterprise group with a large firm). There is a risk that those firms with a large parent company may differ in systematic ways from those without (they may, for example, face less acute constraints in financial markets), though some mitigation might be found by focusing on the subset of SMEs with large foreign parent companies who do not qualify under ESOS. However, findings under this option could not be generalised beyond SMEs.
- **Notifiers compared to non-notifiers:** Comparisons made between those who had and had not notified compliance at a previous set point in time could offer more generalisable findings regarding the impacts of ESOS. However, differences between these two groups are also likely to be systematic (for instance, those who had delayed their notification might be thought to give lower priority to energy efficiency issues, biasing comparisons between the two groups).

Analytical Methods

The issues created by the choice of counterfactual can potentially be addressed through the application of appropriate analytical techniques. The following range of quasi-experimental approaches were considered (alongside non-statistical approaches which are described in more detail below):

A: Tracking notifiers before and after the policy (with no involvement of a comparison group)

B: Comparing obligated versus non-obligated (matching with longitudinal panel analysis)

C: Comparing early versus late notifiers (pipeline design with no comparison group)

D: Comparing notifiers versus non-notifiers (based on status at a previous given point in time) using Instrumental Variables

E: Comparing those just in and just out of scope of the policy (Regression Discontinuity Design)

These approaches were assessed according to the following criteria:

- Robustness[8];
- Analytical confidence issues
- Implementation risk;
- Generalisability of findings to ESOS population; and
- Data availability.

The relative weighting given to these criteria will ultimately be critical in selecting an impact evaluation method that best meets the needs for the evaluation, the type of evidence that is required and the key questions it seeks to answer (an emphasis on accountability over learning, or vice versa, could affect the direction of the evaluation, for example).

Overall, and based on an equal weighting of the appraisal criteria, this assessment did not suggest a dominantly preferred option. A mixed-evaluation strategy is therefore recommended that makes use of non-statistical approaches to exploring ESOS' impact as well as comparison groups that include non-obligated organisations and any organisations who had not notified compliance.

The appraisal of each potential method demonstrated that more robust approaches were either infeasible or offered low generalisability. Approaches offering more generalisable findings tended to be less robust and carried significant analytical hazards or implementation risks. Overall, the following options were the highest scoring across the appraisal criteria (if given equal weight):

B: Comparisons using non-obligated SMEs: Propensity Score Matching (or equivalent) matching could be used to identify a sample of non-obligated SMEs as a counterfactual that share similar features to obligated SMEs. Longitudinal panel methods (difference-in-differences or fixed effects) could be applied to provide results that are robust to any time invariant differences between groups.

D: Notifiers versus non-notifiers using Instrumental Variables: The likelihood of notifying compliance to meet the deadline was connected to the density of lead assessors located in proximity to organisations (likely partly related to the lead assessor's role in raising awareness of the scheme). If the location decisions of lead assessors can be assumed to be independent of the organisational characteristics determining energy efficiency behaviour, then it may also be possible to obtain

[8] Using the Maryland Scale of Scientific Methods as a guide

robust results using an instrumental variable strategy. However, this assumption cannot be tested, creating analytical hazards for this approach.

To compensate for these issues, it is proposed that non-statistical approaches could be integrated into the design of the evaluation. An approach grounded in realist evaluation principles, for example, would help develop an in-depth understanding of what elements of the ESOS policy do and do not work in achieving its intended objectives, for whom and under what circumstances. However, non-statistical approaches carry the disadvantage that the scale of the effects involved cannot be determined, preventing a cost-benefit analysis.

Economic evaluation

An economic evaluation of ESOS could seek to involve two key elements: cost-effectiveness analysis, exploring the unit cost of the impacts delivered relative to other similar initiatives; and, cost-benefit analysis, placing a monetary value on the impacts delivered and relate these to the costs involved. Data would be available which measured the costs and benefits of the scheme, though the feasibility of an economic evaluation will be contingent on the extent to which a quantitative impact approach would generate robust estimates.

Implementation strategy

The implementation strategy for a longer-term evaluation of ESOS, based around the considerations and recommendations outlined above, will need to involve the following key workstrands, data collection approaches and audiences.

- **Stakeholder engagement:** Obligated organisations, policy stakeholders, scheme management teams, assessor market representatives and external stakeholders (such as, energy efficiency technology manufacturers and installers) will all be important audiences to engage in the impact evaluation. Gathering a wide range of external perspectives on the achievements of the scheme, and in the context of wider factors, will be particularly crucial for non-statistical approaches which rely on the collection of rich qualitative data to a large extent.

- **Primary research:** An impact evaluation will require the following primary data collection approaches to be applied to engagement with these audiences: quantitative survey and in-depth qualitative case-studies with obligated organisations; and, in-depth qualitative interviews with the full range of internal and external stakeholders.

- **Datalinking:** Two main strands of data linking could also form part of the mainstage specification: data linking to energy use data, and data linking to firm-level productivity and expenditure data. The timing of the analysis will need to take account of lags in the publication of data. It is anticipated that baseline data for 2014/2015 is likely to be available in 2017/2018 while outcomes data for 2017/2018 is likely to be available by around 2020. This may mean a separate report based on this data may be required following reporting of the primary evidence described above (which can be completed in 2018).

A key risk facing the implementation of a full impact evaluation of ESOS is the quality of data on which analysis is being performed and conclusions drawn. Current challenges to data quality are a result of the ongoing work by the scheme compliance bodies creating a lack of certainty on the eventual nature and scale of the obligated population and company hierarchies. While some assumptions can be made, this uncertainty may affect confidence in the assignment of treatment and therefore would pose a significant risk for the detection of causal effects.

1.0 Introduction

Ipsos MORI and University College London (UCL) were commissioned by the Department of Energy and Climate Change (DECC)[9] in April 2015 to undertake a process evaluation of the Energy Savings Opportunity Scheme (ESOS) and to conduct a feasibility study into the design of a full impact evaluation. The findings of the feasibility study are presented in this report[10].

1.1 Overview of ESOS

Article 8 (4-6) of the EU Energy Efficiency Directive (2012/27/EU) requires the UK to mandate that all large undertakings (and smaller organisations part of a group including a large undertaking) undertake audits or a specified equivalent (such as ISO 50001) of the energy used by their buildings, industrial processes and transport to identify cost-effective energy saving measures by 5th December 2015 and at least every four years thereafter. ESOS was developed by DECC to implement the Directive in the UK. Obligated organisations were required to notify compliance with the scheme administrator - the Environment Agency (EA)[11], and implementation of the energy saving measures identified was voluntary.

1.2 Impact evaluation aims and approach

Six high-level questions were set by DECC overall for the evaluation of ESOS:

KQ1: How have large organisations reacted to ESOS?

KQ2: How has the assessor market responded?

[9] The Department for Business, Energy and Industrial Strategy (BEIS) as of July 2016

[10] A separate report presenting process findings, and exploring early indications of impact, is available here: https://www.gov.uk/government/publications/energy-savings-opportunity-scheme-esos-evaluation-of-the-scheme

[11] The Environment Agency is the scheme administrator for the whole of the UK. Responsibility for compliance and enforcement rests with the Environment Agency in England and the equivalent devolved agencies i.e. Scottish Environment Protection Agency, Northern Ireland Environment Agency, Natural Resources Wales; and the Secretary of State for Business Energy and Industrial Strategy for organisations with mainly offshore activities.

KQ3: How is ESOS influencing organisational energy efficiency policy and practice?

KQ4: Has ESOS been implemented in a way that has avoided unnecessary burden and cost?

KQ5: What impact has ESOS had on organisations?

KQ6: How have the benefits of energy efficiency been realised by large undertakings?

The overall aim of the scoping study presented in this report was to assess the feasibility of, and most appropriate approach to, conducting a longer-term impact and economic evaluation of ESOS which answers key evaluation questions four to six. The feasibility study has focused on addressing two central challenges:

- How far it is possible to establish a suitable comparison group to support an impact evaluation?
- What analytical methods could potentially be applied in such a study and what limitations, if any, do they have in answering the key evaluation questions?

These questions have been explored through the following activities:

- **Evaluation Framework development**, including a descriptive Theory of Change and logic model mapping, to identify the key outcomes and impacts anticipated from ESOS that will need to be identified and monitored through the impact evaluation.
- **Scheme management information analysis,** to understand the key characteristics of ESOS-obligated organisations and to gather information on the likely sample sizes (and therefore the statistical precision) of any later econometric analyses.
- **Review of secondary datasets** that could potentially feed into the measurement of ESOS outcomes in an impact evaluation, considering any limits to the data in terms of access rights, data availability or timing, as well as match rates through feasibility testing of data linking approaches.
- **Appraisal of potential impact evaluation methods** based on this data availability, including quasi-experimental methods at various levels of robustness as well as more theory-based approaches.

1.3 Limitations to the impact evaluation scoping study

The following limitations should be borne in mind when reviewing the feasibility study findings:

- **Ongoing work to bring organisations into compliance** by the scheme administrator means the size of the obligated ESOS population is not yet confirmed, creating some uncertainty around future analytical sample sizes and thus the robustness of some potential impact evaluation approaches.
- **Complex company hierarchies** are discussed within this report, with implications for understanding potential counterfactual groups and identifying organisations for datalinking.

2.0 Evaluation Framework

This chapter sets out an overarching framework for the impact evaluation of ESOS. The framework specifies the policy objectives for the scheme, the rationale for intervention, and through defining the 'theory of change' provides a set of key outcomes that will need to be explored through the impact evaluation. It also considers the contextual factors that may influence these outcomes beyond the control of the scheme.

2.1 Energy Savings Opportunity Scheme (ESOS)

The European Union (EU) has a target to reduce primary energy consumption by 20 percent by 2020 (against a business-as-usual projection made in 2007) through improvements in energy efficiency. In 2011, the European Commission estimated that the EU was half-way towards this target based on existing policies. The EU Energy Efficiency Directive (2012/27/EU)[12] was introduced as part of a drive to establish a common framework of measures to promote energy efficiency across different sectors of the economy throughout the EU. The ESOS Regulations 2014 were implemented in the UK in response to Article 8 (4-6) of this Directive. The objectives of ESOS, as set out in the Impact Assessment[13], are to:

- provide large enterprises with enterprise-specific information about how they can make energy savings;
- stimulate the take-up of cost-effective energy efficiency measures;
- minimise the cost to businesses of complying with the regulations, and;
- maximise the synergies with existing policies.

[12] Eur-lex.europa.eu, 2015. Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1399375464230&uri=CELEX:32012L0027

[13] Impact Assessment: Energy Savings Opportunity Scheme, 2014. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323116/ESOS_Impact_Assessment_FINAL.pdf

2.2 Rationale for intervention

The UK Energy Efficiency Strategy (2012)[14] highlighted that 196 TWh of socially cost-effective savings in 2020 could be achieved from energy efficiency investments. It stated that if this potential could be realised, final energy consumption in 2020 could be 11 percent lower than the business as usual baseline. The Strategy highlighted that the commercial and industrial sectors had the potential to contribute some 69 TWh of these savings. In principle, firms have an incentive to implement energy efficiency projects where the expected cost savings exceed the investment costs[15]. However, the strategy identified four areas where market failures and other barriers will inhibit the extent to which firms will implement these potentially profitable investments without intervention from the public sector. The following are of relevance to ESOS:

- **Information asymmetries**: Senior managers may not always have a developed understanding of the costs, benefits, or risks associated with energy efficiency investments[16] [17]. External expertise may be needed to support effective decision making. However, while external experts offer a stronger understanding of costs and benefits involved, they may also be perceived to have an incentive to understate risk or overstate potential cost savings (particularly if they anticipate benefiting from future implementation contracts). These perceptions could create mistrust in recommendations made, reducing willingness to pay for advice and reducing the financial viability of higher cost and quality advisory services (and resulting in suboptimal identification and implementation of profitable energy efficiency schemes). These problems can be overcome by employing internal energy efficiency experts, though this would not be an efficient option for all.

- **Cultural barriers**: Cultural factors - though not strictly a market failure - may act as a barrier to privately and socially optimal levels of investment in energy savings projects. For example, board-level awareness has been highlighted as crucial but challenging to

[14] UK Energy Efficiency Strategy: The Energy Efficiency Opportunity in the UK, 2012. Available at: https://www.gov.uk/government/publications/energy-efficiency-strategy-2013-update (Accessed July, 2015). Although more up-to-date evidence is now available from sources such as the Building Energy Efficiency Survey, this reference is made here as it was a key source during the policy development phase pre-implementation of ESOS.

[15] This is not equivalent to the measure of cost-effectiveness adopted in the UK Energy Efficiency Strategy, which also considered the social value of non-traded carbon emissions and other environmental externalities.

[16] Shining a Light: Uncovering the Business Energy Efficiency Opportunity, 2013. Available at: http://www.cbi.org.uk/media/934998/shining_a_light__uncovering_the_business_energy_efficiency_opportunity_cbi_report_aug_2013_screen.pdf (accessed July, 2015)

[17] Carbon Trust, The Business of Energy Efficiency, 2010

achieve. This is due to a lack of both awareness and necessary skills to implement improvements regarding managing energy use. The issue is made more acute by a lack of joined-up working across and within firms[18].

Key elements of ESOS have been designed to respond to these market failures and barriers. The mandatory nature of the obligations forces firms to acquire information on the costs and benefits of the available investments and behavioural measures every four years. Those providing this advice (including internal staff) must receive external accreditation, potentially offering important quality signals and raising trust. Any energy savings audits produced require sign-off by a board-level director, giving greater visibility at a strategic level on the commercial opportunities offered by energy efficiency investments.

In June 2014, DECC developed an Impact Assessment for ESOS considering the resource costs and potential benefits of the obligation. Based on relatively conservative assumptions around how far ESOS would produce behavioural change amongst large undertakings, it was estimated that large enterprises would benefit in the form of energy savings of around £2.2bn. In addition, projected improvements in air quality were valued at £320m, with the value of future abatement costs associated with reduced greenhouse gas (GHG) emissions valued at a total of £280m[19].

2.3 Theory of Change Framework for ESOS

This section describes the expected causal process by which ESOS will deliver its intended results. It describes the inputs and activities involved in the delivery of the scheme and their links to its intended outputs, outcomes and impacts. The purpose of the framework is to define the key measures that will need to be monitored through an impact evaluation of the regulation.

2.3.1 Inputs

The ESOS Impact Assessment presented the following expected costs for the scheme under the chosen option of mandatory notification of compliance but with the option for businesses to voluntarily disclose further information (Policy Option 2 in the Impact

[18] Shining a Light: Uncovering the Business Energy Efficiency Opportunity, 2013. Available at: http://www.cbi.org.uk/media/934998/shining_a_light__uncovering_the_business_energy_efficiency_opportunity_cbi_report_aug_2013_screen.pdf (Accessed July, 2015)

[19] Impact Assessment: Energy Savings Opportunity Scheme, Department for Energy and Climate Change and Department for Transport, 2014. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323116/ESOS_Impact_Assessment_FINAL.pdf (accessed July 2015).

Assessment). Costs incurred as a later result of ESOS compliance (e.g. investment costs) are treated as an outcome of the policy rather than as an input.

Figure 2.1: Costs incurred in delivery and compliance with ESOS

Cost	Incurred by	Estimated present value over 16 years[20]
Cost of audits	Businesses in scope	£165m
Additional administrative burden (e.g. accompanying auditors, compiling Evidence Packs)	Businesses in scope	£235m
Auditor accreditation	Auditor professionals	£20m
Scheme administration	Government	£10m

Source: DECC Impact Assessment[21]

2.3.2 Activities

ESOS is a regulatory intervention, and most activities associated with the scheme relate to securing and monitoring compliance with those regulations. Delivery of the regulation involved the following activities.

Awareness raising activity: DECC and the Environment Agency implemented awareness-raising activities to give relevant organisations information on the scheme's requirements and to establish their eligibility under ESOS. This included direct mails, roadshows, a scheme guidance publication and helpdesk service. A parallel process evaluation has found that channels external to the scheme have also been important elements of this activity, such as industry press and assessor marketing.

Establishing registers of accredited ESOS Lead Assessors: The establishment of multiple registers of accredited assessors is a central element of the scheme providing quality signals needed to address the issues associated with asymmetric information (as

[20] The present value of costs (discounted at the Green Book rate of 3.5 percent) over 16 years (i.e. four ESOS cycles)

[21] ESOS Impact Assessment, 2014. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323116/ESOS_Impact_Assessment_FINAL.pdf (Accessed July, 2015)

well as ensure that there are sufficient levels of competition between suppliers of the advisory services required). Accreditation of Lead Assessors was externalised to professional bodies approved by the Environment Agency. To secure approval, relevant organisations were asked to complete and submit a pro-forma containing evidence of how they guaranteed quality throughout their members. Fifteen registers of Lead Assessors were approved, some of which existed prior to ESOS and included assessors who undertake audits as part of other, related, non-domestic energy efficiency schemes (e.g. for the CRC Energy Efficiency Scheme) whereas others were set up solely to approve individuals to be ESOS Lead Assessors.

Guidance for Lead Assessors: No 'official' guidance has been produced by BEIS or the Environment Agency specifically for Lead Assessors. This choice was made to give flexibility regarding the choice of auditing and sampling approaches and to avoid duplication of energy auditing standards and approaches already available. The question of how far this approach minimised costs incurred by obligated parties should be explored in the evaluation. If there is less than full clarity, a risk-averse firm or assessor may over-specify requirements to minimise the risk of non-compliance. On the other hand, cost-minimising firms may prefer to commission low quality audits (leading to low trust in the recommendations involved).

Guidance for Obligated Parties: The Environment Agency produced a guidance document outlining the requirements of the Scheme for qualifying organisations[22] and provided a dedicated e-mail 'helpdesk' to field queries. Other business advice and support organisations (e.g. Resource Efficient Scotland) and devolved administrations also took on an advocacy role and fielded questions. In addition, DECC published an 'Approaches to ESOS' audit guidance[23].

External energy audits: One route for compliance is the procurement and delivery of an externally procured or internally delivered audit. ESOS audit guidance recommends at least one site visit but makes no other stipulations about the audit process in terms of total number of site visits, which areas of energy use are included (10% of an organisation's total energy use does not have to be covered under ESOS regulations), and the format of the resulting audit report. Audits were anticipated to cover:

- financial savings offered by the available improvement measures;

[22] Complying with the Energy Savings Opportunity Scheme (ESOS), 2015. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/445205/LIT_10094.pdf (Accessed July, 2015)

[23] Approaches to ESOS Audits, 2015. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/435845/DECC_Approaches_to_ESOS_audits_guide_FINAL_20_03_15_amended.pdf (Accessed July, 2015)

- return on investment measures (and other economic criteria agreed);
- possible non-energy gains (such as productivity or maintenance savings),
- comparisons between alternative energy efficiency improvement measures; and
- technical interactions between multiple actions[24].

Compliance requires sign off on the audit report by a senior board director (and those using an in-house assessor must have the audit report signed off by two senior board members) who are required by law to have 'seen and considered the recommendations' provided by the audit or alternative compliance routes. Organisations should also compile an Evidence Pack containing evidence that the audit took place, the energy consumption identified during the audit, and evidence that the report has been signed off by at least one board level director within the organisation and a Lead Assessor.

ISO 50001: Organisations with an ISO 50001 energy management system certified by an accredited certification body (and covering all energy use across the corporate group in the UK) were considered ESOS compliant without any requirement for further auditing or appointment of an ESOS Lead Assessor. This route to compliance still required board director sign-off on ESOS compliance notification information submitted and confirmation that they had reviewed findings of the ISO 50001 certification.

Notification, Compliance and Quality Assurance Checks, and Publication: Organisations were obligated to notify the Environment Agency of their compliance by the 5th December 2015. The notification process required organisations to provide basic information, such as whether they used an internal or external assessor, whether they represent a disaggregated firm complying separately to a larger parent, and details of the senior board-level member providing sign-off. The notification process did not require submission of information about the audit.

The Environment Agency is carrying out investigations to check compliance notices submitted (and this process is still ongoing). Checks are largely desk-based, and where it is suspected that incomplete, or inaccurate information has been provided, organisations may be required to submit their Evidence Packs and proof of director level sign-off to the Environment Agency. The Environment Agency has also conducted 209 audits of compliance notices as at August 2017.

[24] Approaches to ESOS Audits, 2015. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/435845/DECC_Approaches_to_ESOS_audits_guide_FINAL_20_03_15_amended.pdf (Accessed July, 2015)

Enforcement: The Environment Agency's guidance around ESOS states that obligated organisations that do not meet the requirements may be liable to compliance and enforcement activities. These include a compliance notice (an information request), an enforcement notice telling the organisation what they must do to comply with a requirement of ESOS, or a penalty notice involving civil penalties (up to £50,000) for breaches of ESOS regulations.

2.3.3 Outputs

Awareness and understanding of ESOS and its requirements: As noted above, marketing and communications activities were critical in raising the awareness of ESOS amongst large firms needed to meet compliance and broader policy objectives.

Decision to Comply: Once an organisation becomes aware of their obligations, it is hypothesised in this Theory of Change that a decision will be made regarding whether to pursue action to comply with the scheme. The decision of a purely profit-maximising firm would involve a trade-off between the anticipated costs and benefits of compliance and non-compliance. This assessment will be influenced by the costs and perceived benefits associated with compliance, the threat of detection for non-compliance and enforcement of penalties, and the nature of those penalties (including reputational costs). While the decision to comply with ESOS may be a purely economic decision for some firms, others may be encouraged to comply for reputational reasons or to align with their environmental or sustainability agenda and goals.

As at 30 June 2016, the EA had received notifications from 81 organisations that were not obligated under the regulations but notified compliance 'to show commitment to saving energy'. There were some anecdotal suggestions from stakeholders engaged through the scoping process that SMEs saw compliance as beneficial in supporting transition, where growth is expected that will lead them to be eligible in the future, or because they sought reputational benefits from being publicly listed as a compliant organisation.

Allocation of responsibility for ESOS compliance within organisations: Evidence gathered from stakeholders through the scoping phase suggested that the impact of ESOS may be linked to decisions made by obligated organisations on where responsibility for compliance sat. It is anticipated that this will differ between organisations: some may see compliance as a sustainability issue, whereas for others it might sit within the facilities or compliance departments. These internal teams may have greater or lesser scope or incentive to act on the recommendations of an ESOS audit, and the process evaluation will need to explore how these decisions were reached by obligated parties.

Selection of Compliance Route: An obligated organisation will look to select the optimal route for compliance. A hypothetical profit-maximising firm would adopt the route that minimised the net-cost of securing compliance. This will be influenced by the relative costs

of compliance via ISO 50001, internal accreditation, procurement of external advisory services, the complexity of organisations to be audited, level detail desired, and the benefits that are expected to arise. Given the market failure issues identified above, there is a risk of a self-fulfilling mechanism occurring if organisations seek out the lowest cost option as a result of low trust in the quality of the audit (with those views confirmed as a consequence of the choices made). Some stakeholders engaged suggested ESOS may also some large organisations to externalise existing internal energy management functions (potentially working against the underlying objectives of the regulation).

2.3.4 Outcomes

Implementation of Energy Efficiency Measures: The scheme purposefully places no compulsion on organisations to act on ESOS audits. However, in principle, a profit maximising firm has an incentive to implement all recommendations where the present value of expected returns exceeds those offered by risk-free assets (assuming perfect financial markets and no transaction costs).

Evaluations of previous energy audit schemes have gone some way to demonstrating the extent to which energy efficiency recommendations are implemented in practice. Analysis of the Carbon Trust's energy efficiency audits with large businesses[25] showed audits identified measures with an average annual saving of over £200,000, or around 4000 MWh (with over 70 percent of these involving a payback period of fewer than three years). However, only 30 percent of recommendations went on to be implemented by businesses, suggesting that some but not all profitable investments are made in practice. This research also showed that, on average, organisations can take between two and four years to implement energy efficiency recommendations. Uptake could also be limited by:

- **Organisational factors**: Uptake could be affected by the budgets available to - and the capacity and skills of - staff to put the recommendations into practice. Implementation of recommendations also requires the responsible individual within the organisation (likely to be an energy or facilities manager) to be a change agent for that organisation. This means being well networked, influential and able to bring about recommended improvements where there is a financial incentive to do so.

- **Competing investments**: Where energy efficiency investments are considered alongside other proposals, it may be considered by decision makers that there may be more profitable areas of focus for the firm if the firm is budget constrained.

[25] Energy Savings from Audits: Analysis from the Carbon Trust's Close Out database.2015. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323113/ESOS__Analysis_of_the_Potential_for_Energy_Savings_from_Audits_FINAL.pdf (Accessed July, 2015)

- **Quality of audit information and level of belief or trust in recommendations**: The willingness of an organisation to act on the recommendations made by the lead assessor will depend on the perceived quality of the audit and level of trust in the recommendations made (as stressed above).
- **Business continuity**: Uptake of energy efficiency measures can be affected by business continuity issues e.g. requirement for server downtime – although, in theory, these disruptions will be priced into the cost of implementing the recommendation.
- **Perceived rate of technological change**: If an organisation believes the rate of technological progress with energy efficiency measures is high then they may be encouraged to delay action until a time at which they think a new, more efficient technology will be available at a lower cost. It will be important to understand the basis on which organisations make investment decisions and the extent this is a factor in that process.
- **Energy Efficiency Investment**: Implementation of energy efficiency measures will also be accompanied by an increase in energy efficiency investment (i.e. capital expenditure). This may not apply in the case of implementation of behavioural measures.
- **Effects in Secondary Markets**: In addition to its effects on organisations required to comply with the scheme, ESOS will also have effects in secondary markets:

i) Assessor market

The advisory services required to demonstrate compliance with obligations under ESOS were generally provided by the private sector. Assessors had an incentive to enter the ESOS market if they anticipated profits in excess of the revenues associated with undertaking other activities (a calculation which was in part dependent on the costs associated with obtaining accreditation). The strength of incentives may potentially be limited by the four year cycle of ESOS audits. However, some may seek to develop an on-going service which extends over the four year period, which could include contracts to deliver implementation services, including such as financed energy conservation measures (ECMs) which might include guarantees through energy performance contracts (EnPCs) – though it is important to note that these motivations drive some of the market failures identified above.

ii) Technology or service innovation using ESOS data

It is possible that assessor organisations undertaking large volumes of ESOS audits would generate large scale datasets with regard to the specific energy efficiency needs of large organisations and firms. This could in turn drive service or technology innovation, allowing the assessor organisation to provide services more cost-effectively. In turn, this could raise demand for these services, leading to further investment in energy efficiency measures

and further carbon dioxide (CO2) savings. The process evaluation will need to explore how far any effects of this nature have been achieved (though they are of secondary importance, given the objectives of the scheme).

iii) Technology Producers

Finally, to the extent that ESOS stimulates an increase in demand for energy efficiency technologies, there may be benefits for producers of those technologies (which might be explored as part of a cost-benefit analysis). However, in line with the general principles set out in the HM Treasury Green Book Appraisal Guidance[26], an assumption has been made that any increase in demand will place pressure on the price of key inputs (e.g. labour, raw materials, property) leading to an offsetting reduction in demand elsewhere in the economy.

2.3.5 Impacts

Firm level impacts

Assuming ESOS leverages changes in organisational behaviour (either in the form of investment and installation of new capital items or in the form of behaviour change) that would not have occurred in its absence, a range of broader effects might be observed amongst those obligated under ESOS in the longer term.

BEIS analysis estimates that ESOS will reduce energy consumption by 3TWh each year, with businesses saving over £250m on their energy bills and that the scheme will deliver a net benefit to the UK of £1.6 billion NPV (calculated over 15 years), through the following effects:

- **Reduced energy consumption**: The implementation of energy efficiency measures will reduce consumption of energy by the obligated organisation and leading to a reduction in energy costs for a given level of output or volume of sales.
- **Reduced maintenance expenditure**: The installation of better performing energy management systems, or specific energy efficiency measures (for example longer lasting LED bulbs), may lead to lower maintenance spend.
- **Overall productivity**: The firm will likely see an increase in productivity through a reduction in unit costs. However, the size of this effect will depend on how far energy efficiency investments have crowded out other investments with the potential to raise productivity. For example, if an organisation is encouraged to invest in energy saving measures due to the ESOS audit, investment may be diverted away from other capital

[26] https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent

projects (if it faces capital constraints). As such, it is anticipated that an impact evaluation will need to determine effects on both total capital investment and investment in energy efficiency measures to explore the full range of causal effects of interest.

- **Increase in output**: Firms may also choose to reduce the price of its output, enabling it to expand its market share and encouraging it to expand overall production. In order to meet these higher levels of demand (and realise the revenue benefits), energy consumption may also expand, offsetting initial reductions in energy consumption. This is known as the 're-bound' effect, and highlights the need to examine both energy efficiency (e.g. energy consumption per unit of output) alongside total energy consumption and total output (GVA) as part of an eventual impact evaluation.

- **Displacement:** However, while the rebound effect may limit the size of energy consumption reductions amongst organisations exposed to the ESOS obligation, there may be displacement effects in product markets that work in the other direction: where firms reduce their output prices, it can be anticipated that any following increase in market share will be taken from competitors. To the extent these firms are based in the UK, this will lead to improvements in social welfare through a transfer of output from less to more efficient producers. If this is also accompanied by a transfer of output from less to more energy efficient producers, the overall reductions in energy consumption may exceed those driven by the initial improvement in energy efficiency.

2.4 Contextual issues

There are a range of external factors that have the potential to influence the likelihood that ESOS achieves its stated aims and objectives that may need to be accounted for in the design of the impact evaluation.

2.4.1 Wider policy landscape and uncertainties

Industrial energy efficiency is central to the Government's energy policy as the 2011 Carbon Plan made clear[27]. While ESOS has been developed as the Governments' main programme delivering information and advice to larger non-domestic organisations, it operates within a complex policy landscape. Other UK climate change policies (such as the CRC Energy Efficiency Scheme (CRC)) require organisations to measure and manage energy consumption accurately, although the eligibility requirements and scope of energy and emission reporting vary across these mechanisms.

[27] The Carbon Plan - Reducing Greenhouse Gas Emissions, 2013. Available at: https://www.gov.uk/government/publications/the-carbon-plan-reducing-greenhouse-gas-emissions--2 (Accessed July, 2015)

Over the evaluation period there have, however, been changes to the wider landscape with the announcement of reforms to improve the energy tax and reporting regime, including abolishing the CRC and absorbing the price signal into the Climate Change Levy[28]. It was also announced that a new streamlined energy and carbon reporting framework will be introduced by April 2019. This will not obviate the need to control for exposure to these policies in the past: the assumption in the DECC Impact Assessment for ESOS is that where businesses have participated (or have been obligated) under these policies in the past, they will have already partly or fully exhausted opportunities to improve energy efficiency (limiting the potential effectiveness of ESOS in these cases).

It is therefore important that the evaluation of ESOS is situated within a clear understanding of these different policies, and how they interact, as this will affect how different types of companies' view ESOS. There may also be implications for whether organisations are likely to exploit the findings of the ESOS audit given their exposure to other reporting obligations and energy management systems. A short summary of the potential overlap and anticipated impact is provided below:

EU Emissions Trading System: EU ETS compliant companies[29] tend to be large and highly energy intensive. They are likely to have established, sophisticated energy auditing systems in place. ESOS is unlikely to carry serious administrative burdens for these organisations, or add significant value at a process level as all the available savings should have - in principle - been realised. However, ESOS may help with more marginal energy uses such as transport and office buildings or where EU ETS companies have a small number of assets in the scope of the scheme amongst a much wider range of assets and activities that are not.

Climate Change Agreements (CCAs): An operator that has a CCA[30] will measure and report its energy use and carbon emissions against agreed targets over four, two-year target periods. Companies that have agreed CCAs tend to have high energy costs as a

[28] HMRC, Climate Change Levy: main and reduced rates, 2016, https://www.gov.uk/government/publications/climate-change-levy-main-and-reduced-rates/climate-change-levy-main-and-reduced-rates.

[29] The current Phase of the EU ETS, Phase III, runs from 2013 to 2020, and includes approximately 1,100 UK participants (including power stations, oil refineries, offshore platforms, and manufacturing plants producing iron and steel, cement and lime, paper, glass, ceramics and chemicals). Other organisations, including universities and hospitals, may also be covered by the EU ETS depending upon the combustion capacity of equipment at their sites. Aviation operators flying into or from a European airport are also covered by the EU ETS.

[30] The current CCA scheme started in April 2013 and will run until 31 March 2023. CCAs are available for a wide range of industry sectors from major energy-intensive processes such as chemicals, paper and supermarkets to agricultural businesses such as intensive pig and poultry farming.

proportion of turnover but are a more mixed community of large and smaller businesses. Again, the added value of ESOS amongst any large organisations that have entered into CCAs may be limited.

CRC Energy Efficiency Scheme: Under CRC, organisations have to report total energy supplies (actual or estimated) and energy generated by renewables if applicable. By comparison, an ESOS audit requires a calculation of total energy consumption, broken down by different assets and activities in an energy consumption profile as well identifying energy efficiency opportunities and energy saving measures. CRC-compliant companies are a very mixed community[31] with a significant overlap with ESOS (estimated at about 30 percent). The additional effect of ESOS is likely to depend on company type and sector.

The Government is also clear that ESOS needs to be deployed to minimise the cost to business and maximise synergies with the broader suite of industrial energy efficiency policies. The impact evaluation will need to assess how well ESOS has met these objectives to date and determine the added value it has delivered.

2.5 Summary of Key Outcomes of interest

Table 2.1 summarises the key outcomes that were identified in the theory of change and logic model for ESOS. The table also gives a measure of priority (high, medium, or low) based on the centrality of those measures in assessing how far ESOS has addressed its central objectives and indicative timescales over which the effects of compliance with ESOS in the 2014 round might be observed (i.e. the future rounds of compliance are not considered).

In broad terms, it is assumed that the core effects of ESOS on the energy efficiency behaviour of obligated organisations are central to a main-stage evaluation, while demonstrating effects in secondary markets may be less critical. Additionally, it is assumed for this theory of change that all effects on organisational behaviour will be realised before the next ESOS compliance deadline in 2019 (though these effects on behaviour may have longer term effects on energy demand and productivity).

[31] Participants include supermarkets, water companies, banks, local authorities and all central government departments. Inclusion is mandatory on the basis of meeting a threshold for electricity consumption (6000MWh) through half hourly settled meters or mandated for being a public body.

Table 2.1: Key Outcomes

Outcome	Priority	Expected timescale over which effects might be observed
Effects on obligated firms		
Investment in energy efficiency measures (including transport-related)	High	2016-2019
Implementation of behavioural efficiency measures	High	2016-2019
Energy and fuel consumption	High	2017 onwards
GHG emissions	High	2017 onwards
Productivity	High	2017 onwards
Output (GVA)	Low	2017 onwards
Reputation	Medium	2016-2019
Wider environmental effects		
Abatement costs	High	2017 onwards
Air quality	Low	2017 onwards
Treatment costs for morbidity caused by poor air quality	Low	2017 onwards
Effects on suppliers of energy efficiency services and technology		
Sales	Low	2015
Profits	Low	2015
Service and product innovation	Low	2016 onwards

Source: Ipsos MORI

3.0 Measurement of Outcomes

The evaluation framework set out in Chapter 2 highlighted the range of key outcomes that will require measurement during a quantitative impact evaluation of ESOS. This chapter explores the range of mechanisms that might be employed to collect these quantitative measures, and the hazards that may occur, learning from feasibility testing conducted during this study.

3.1 Sources of evidence considered

A robust impact evaluation requires measurements of the key outcomes of interest (as set out in Table 2.1) both before and after the intervention for an appropriate set of observational units (in the case of ESOS, these units are the obligated organisations and other businesses indirectly affected). Such measures can potentially be derived from the following sources:

- **Administrative datasets**: The Government and other organisations in both the private and public sectors collect large amounts of information regarding individual organisations through routine administrative processes. These datasets have a number of core advantages over alternative sources of information. Firstly, they tend to provide close to comprehensive coverage and carry a high degree of accuracy as non-provision of information can carry penalties. As the data is used to plan routine public or private activities, it is refreshed on a regular basis, resulting in the type of longitudinal data that would optimally drive an impact evaluation. Administrative information does carry limitations in terms of the range of variables that are captured. Administrative data sources include Environment Agency scheme data, Companies House data and Non-Domestic NEED.

- **Existing or planned surveys**: Numerous surveys of organisations are completed by the Government on a regular basis that can potentially be exploited in a main-stage evaluation of ESOS. This included, for example during the initial process and early impact evaluation of ESOS, a regular survey of non-SMEs that DECC commissioned during 2014 to track awareness of ESOS and compliance plans. Further surveys such as this may be expected in future to provide evidence that may be of value to the longer-term impact evaluation of ESOS. Sample surveys can be weaker sources of information than administrative datasets as they typically provide less comprehensive coverage of the populations of interest, and in some cases, may be subject to measurement error. The data from the Annual Business Survey is the main example of this type (see section 3.4 for more information about the survey).

- **Primary survey research**: While there are attractions in maximising the use of existing datasets (for reasons of both cost-effectiveness and quality), there may be either gaps in coverage or lags in the availability of such data. Options for filling these gaps or providing more timely evidence through primary surveys of organisations obligated under ESOS are considered further below (in Table 3.1 below and Section 3.5)

Table 3.1 maps out the range of sources considered as part of this exercise and their possible role in measuring the outcomes described in Table 2.1 in the preceding section. A description of each of the main sources and the issues anticipated is provided after the table.

Table 3.1: Mapping of Outcomes to Data Sources (coverage of outcome measure in data source indicated by shading)

	Existing administrative or survey datasets			**Existing surveys**		
Outcome	**EA scheme data**	**Companies House**	**ND-NEED**[32]	**Annual Business Survey**	**Primary Surveys**	**Summary of Issues**
Details of obligated organisations						
Obligated organisations						Complexities identifying obligated organisations from publicly available data (e.g. getting accurate and up-to-date data on private funding sources for education-sector organisations) mean the obligated population size is not certain.
Identifying subsidiaries' compliance status						Complexities in company hierarchy data (and incomplete lists of subsidiaries in compliance forms) mean there are uncertainties identifying individual organisations.
Trading addresses	Head-quarter sites only	Individual trading address				
Outcomes						
Investment in energy efficiency measures				Overall capital investment only		Energy efficiency investments cannot be isolated & while measures would be net of any substitution or crowding out caused by ESOS, there is a risk that any impacts are insufficiently large to be identified.
Implementation of behavioural efficiency measures						

[32] Non-domestic National Energy Efficiency Data-Framework, described fully here: https://www.gov.uk/government/statistics/the-non-domestic-national-energy-efficiency-data-framework-nd-need

	Existing administrative or survey datasets			Existing surveys		
Outcome	**EA scheme data**	**Companies House**	**ND-NEED**[32]	**Annual Business Survey**	**Primary Surveys**	**Summary of Issues**
Energy & fuel consumption				Combined proxy measure based on spending		Linking meters to organisational sites has many analytical hazards. ABS would (& primary data may) require conversion from prices paid to create proxy measure for energy demand.
GHG emissions						Cannot be directly observed but can be approximated based on changes in energy consumption.
Productivity						
Output (GVA)						Could be combined with energy use to provide longitudinal energy efficiency measure (units of energy consumed per unit of output produced).
Reputation					Perception-only based measurement	Not suitable measure for quantification.
Wider environmental effects						
Abatement costs						Cannot be directly observed, but can be approximated on the basis of changes in energy consumption.
Air quality						Cannot be directly observed, but can be approximated on the basis of changes in energy consumption. Robust attribution of environmental (and knock-on health) effects to ESOS is, however, unrealistic.
Treatment costs for morbidity caused by poor air quality						
Effects on suppliers of energy efficiency services and technology						

	Existing administrative or survey datasets			Existing surveys		
Outcome	**EA scheme data**	**Companies House**	**ND-NEED**[32]	**Annual Business Survey**	**Primary Surveys**	**Summary of Issues**
Sales					Estimation from future investment or auditing / energy consultancy planned	The Business Structure Database could also be used to capture turnover effects.
Profits						
Service and product innovation						
Summary of dataset issues			*Incomplete data coverage (~ 50%)*	*No data for fewer than 250 employees*	*Self-reported & potentially low response rate for organisations who have not notified compliance*	

Source: Ipsos MORI

3.2 Scheme Monitoring Information

Data collected through the administration of the scheme by the Environment Agency provides details of both the population of obligated organisations and those that have notified. The database is based on records filed at Companies House (purchased from a commercial aggregator) that identifies all individual enterprises (13,223[33]) within an enterprise group where at least one member of that group meets the minimum turnover or employment thresholds under ESOS. The data provides a range of key details in relation to obligated organisations:

- Companies House Reference Number
- Headquarters location, Ultimate UK parent company & Overseas parent
- Notification status & Date of compliance
- Trading name, Address
- Number of UK organisations included in notification
- Lead assessor body (voluntary)

While most of this information is of high quality and accuracy, there are some weaknesses in the dataset that need to be considered:

- **Individual trading addresses**: The data provides address details of headquarter sites, but not individual trading addresses. To link records of the ESOS population to energy meter records, details of these individual trading addresses are required (creating problems for multi-site firms or organisations).
- **Notification status:** Organisations are required to report the number of active organisational entities covered by their compliance notice but not to systematically list the other enterprises that are covered. As such, there would need to be further analysis to identify which individual organisational entities have been covered by ESOS notifications.

[33] This is based on ESOS population data produced by the Environment Agency in March 2016.

3.3 Energy Consumption Data

Measuring the energy use of organisations obligated under ESOS will be central to understanding how far the policy has achieved its core aims. This section discusses potential approaches for this, with the key data sources and processes involved and findings from feasibility testing summarised below:

ESOS Outcome measure: ENERGY USE	
Data required	Meter-level records of gas and electricity consumption
Potential source(s)	ND-NEED for meter profile data (or raw meter data as alternative)
Process required	Address matching between ESOS-obligated sites and meters
Key challenges/ limitations	Over-coming differences in address formatting. Time-lag in publication of ND-NEED data, as well as limited meter data coverage (less than 50%) & no transport data. Challenges working with raw meter data included identification of domestic records
Key dependencies	Approvals to access raw meter data, if alternative approach taken
Key learning for mitigating challenges from scoping work	Develop bespoke address matching routine tailored to ESOS; match to multiple (temporally aligned) address databases

3.3.1 Outcome data sources considered

The Non-Domestic National Energy Efficiency Data-Framework (ND-NEED) is likely to be the most suitable source of this information as it provides actual annualised electricity and weather-corrected gas consumption figures from meter readings matched to individual sites' addresses[34]. Alternatively, while this meter profile information has advantages (such as distinguishing domestic from non-domestic data), using raw meter data may be preferable given limited data coverage in ND-NEED (approximately 49% for electricity meters and 30% of gas meters[35]).

[34] Other options excluded due to greater weaknesses include: Display Energy Certificates which are mainly for public sector buildings outside the scope of ESOS; and Energy Performance Certificates based on modelled rather than actual consumption.

[35] According to BEIS information from March 2015.

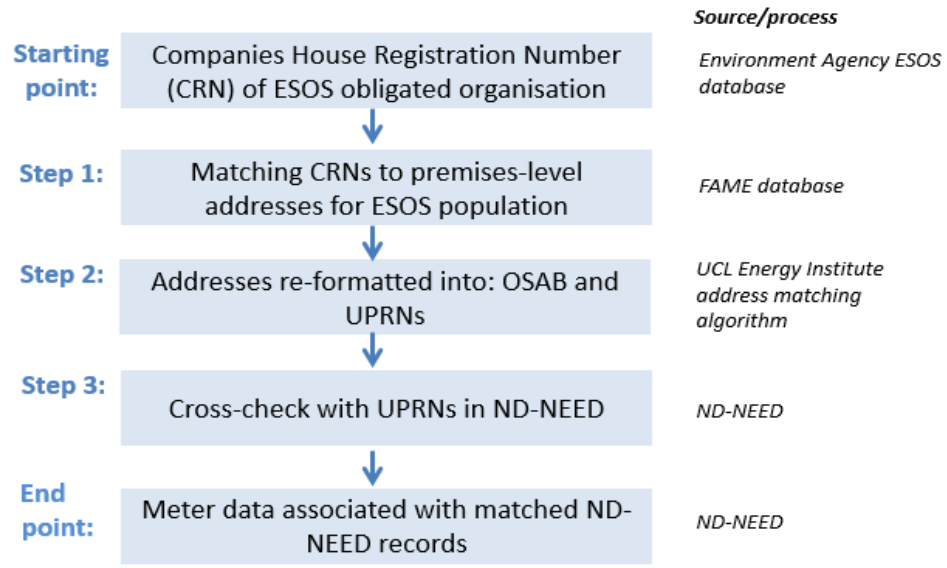
3.3.2 Discussion of scoping work conducted

This section outlines the two matching approaches undertaken in this study: Matching ESOS records to ND-NEED and, as an alternative approach, matching raw meter data to ESOS records.

Matching ESOS-obligated organisations to ND-NEED records

The matching of ESOS-obligated organisations to meter data requires a series of data linking stages, as shown in Figure 3.1 and described further below. Overall, while the feasibility testing of this process showed promise in the ability to identify ESOS-obligated organisations within the ND-NEED, the limited data coverage of meter profiles restricts the ultimate scale of matching through to energy data.

Figure 3.1 Data linking stages to match ESOS organisations to energy use data



Source: Ipsos MORI & UCL

Step 1: Matching CRN to premises-level addresses for ESOS population

In order to identify gas and electricity meters associated with each ESOS-obligated organisation, it is necessary to first identify the premises-level addresses for those obligated organisations. Records of the population of obligated parties were re-matched to Companies House records to obtain records of trading addresses[36].

[36] Companies House filings record the trading addresses associated with individual enterprises, and are compiled in commercial databases such as FAME, Amadeus and Experian.

Findings from feasibility testing on Step 1: CRN matches linked headquarters or main site addresses[37] to 96% of ESOS-obligated organisations recorded. A matching exercise conducted to obtain all trading addresses for ESOS-obligated organisations identified 113,000 separate addresses, but encountered difficulties in bulk downloads of address details (to then be fed into the later data linking steps described in this section). Investigations into the feasibility of obtaining high quality premises-level addresses (beyond headquarter or main registered trading addresses) is recommended as further preparatory work under the full impact evaluation.

Key learning: This was a relatively straightforward process for the feasibility work conducted. However, expanding this process to incorporate all trading addresses may bring challenges.

Step 2: Addresses re-formatted to provide OSAB UPRNs

An address matching process is needed to link ESOS-obligated addresses to Ordnance Survey Address Base (OSAB) Unique Property Reference Numbers (UPRNs). This is the main address matching link into ND-NEED. A code was specifically developed for running this process with ESOS-related addresses. It applied the following logic check to records where the matched Levenshtein score was not perfect (1.0)[38]:

- Outer postcodes were checked between the ESOS postcode and the matched postcode. If they did not match then the match was flagged as bad and recorded as an 'Outer postcode mismatch'. This was based on experience that while postcodes are often entered incorrectly, normally the outer postcode will be correct.

- An individual Levenshtein test was run against all organisation names associated with the chosen OSAB address. If the best of these scores was less than 0.75 then the match was flagged as bad and recorded as 'Organisation levenR is low'.

Findings from feasibility testing on Step 2: 57% of ESOS-obligated headquarters/ main site addresses obtained at Step 1 achieved a trusted match to an OSAB UPRN at Step 2. Two sources of ESOS-obligated addresses were tested during Step 2: 1) those generated from Step 1 above, 2) those provided through self-completion fields in the ESOS compliance notification form. As shown in Table 3.2, the former generated the higher match rate to UPRNs and is the recommended approach for the full impact evaluation.

[37] A full implementation of this process during the mainstage Impact Evaluation would need to extend this to all trading addresses identified through the Bureau Van Dijk FAME dataset.

[38] The Levenshtein score provides a measure of the quality of the address match. The Levenshtein distance for a perfect match is 0. The Levenshtein ratio is equal to 1 minus the Levenshtein distance divided by the length of the longer string; thus the ratio for a perfect match is 1.

Table 3.2: Match rates

Address source:	Sample size run[39]	Matches achieved	Trusted match	Avg. Levenshtein ratio	Avg. Levenshtein distance
CRN matched address	2,037	80%	57%	0.89	7.92
ESOS compliance notification form	1,575	74%	48%	0.83	11.81

Source: Ipsos MORI & UCL

Key learning: Further actions recommended to improve the match rate achieved at the full impact evaluation are: Creation of a code allowing matching to three datasets (OSAB, VOA and Experian, for example), with further small improvements possible by temporally aligning the matching datasets (using a version of OSAB that closely matches the date the ESOS addresses are time-stamped for example[40]); and matching to all possible variants of the business name. Further improvements to the match rate may be possible through manual checks and adjustments, though these are likely to be time intensive on the full population sample.

Step 3: Cross-check of UPRNs in ND-NEED

A database query matches UPRNs generated at Step 2 into UPRNs held in ND-NEED in order to gain access to the meter data (provided ND-NEED has made a successful match to the meters).

Findings from feasibility testing on Step 3: Around 19% of the ESOS-obligated UPRNs could be matched to electricity consumption records held in ND-NEED and around 10% to gas consumption records. This is lower than hoped for this final stage in the process and will limit the potential for longitudinal analysis of this outcome variable.

Alternative approach: Matching raw meter data to ESOS-obligated organisational sites

This scoping study has also tested the feasibility of taking an alternative method to tracking energy use as an outcome variable to establish whether any improvement can be made on the findings presented above. This alternative approach has involved matching meter level data to postcodes identified at ESOS-obligated sites. The key steps involved in this alternative process were as follows:

Step 1: Lists of postcodes were produced from sites identified to be in scope of ESOS;

[39] If there is no bias in the types of obligated organisations achieving a match and a similar match rate can be achieved on the full ESOS population (a fair assumption given a random sample was tested).

40 The feasibility testing has been based on 2012 / 2014 data for VOA and OSAB.

Step 2: Meter records were provided (following approvals being provided by BEIS) for all meters found at these postcodes;

Step 3: Meter records were run through a matching system designed by UCL's Energy Institute to identify which meters belong to ESOS sites. This deals with complexities of matching meters to premises by using GIS and the geometry of the building and site to aggregate to Self-Contained Units (SCU).

Findings from feasibility testing of this alternative approach: Based on testing a small sample of raw meter data, 50% of the electricity records and 35% of gas records[41] were able to be matched to postcodes of ESOS sites. While this suggests some potential to improve on the match rates achieved through linking into ND-NEED, these match rates are still lower than hoped for through this alternative approach.

Key learning: In order to maximise the coverage of energy use data across the ESOS obligated population, it may be possible to complement ND-NEED with other meter information sources, such as raw meter data held by BEIS. Overall this testing suggests that whilst it is likely to be possible to match some energy consumption data to most ESOS obligated organisations, this will not be complete at a site level (though this should be able to be controlled for in analysis as it will be known which sites there are and are not data points for). There are also further alternatives to source proxy data, for example, through the ABS which records energy spend data for large organisations.

3.4 Sources of firm productivity and expenditure data

Beyond the key outcome of energy use discussed above, it will also be important for an impact evaluation of ESOS to seek to understand other firm-level effects of the scheme, such as any impact on the productivity of organisations. This section discusses potential approaches to this, with the key data sources and processes involved and findings from feasibility testing summarised below:

[41] This has not been adjusted to exclude postcodes with no access to gas, so this match rate may increase.

ESOS Outcome measure: Firm-level effects (capital investment, energy use, productivity)	
Data required	Firm-level financial data showing investment, expenditure, sales
Potential source(s)	ABS (& Structure Database as supplement for SME data)
Process required	CRN matching between ESOS-obligated sites and ABS microdata
Key challenges/ limitations	Temporal lags in data availability, no SME data in ABS limiting review of firms employing assessors
Key dependencies	Approvals (and associated timelines) to have data linking completed by ONS, access VML & have analysis signed-off for export
Key learning for mitigating challenges	No significant challenges, but additional stages could include linking survey responses to this data

3.4.1 Outcome data sources considered

The Annual Business Survey (ABS) is an annual survey of 62,000 businesses in the UK undertaken by the Office for National Statistics. It collects information that is incorporated into the calculation of key national accounting measures (such as GDP) and is mandatory for all large firms (defined as those with 250 employees or more) to complete annually. The ABS provides the potential to provide substantial and relevant longitudinal data on organisations obligated under ESOS (as well as those organisations employing assessors to provide ESOS audits), provided it is possible to identify those organisations within the microdata. Table 3.1 indicated the outcome measures of interest for ESOS that the ABS collects and listed some of the issues for consideration around these measures: capital investment, energy and fuel expenditure, GVA, productivity[42], sales and profit.

3.4.2 Discussion of scoping work conducted

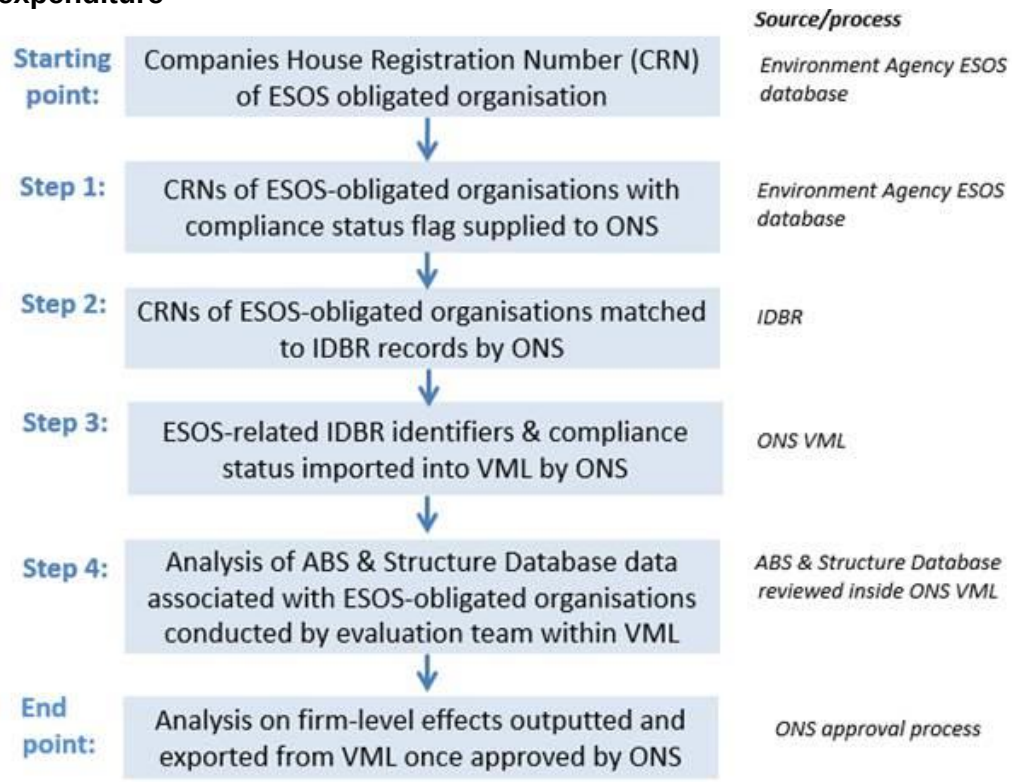
Companies House Reference numbers collected through the Environment Agency scheme data, can be linked to individual firm records in the ABS microdata in the ONS Virtual Microdata Laboratory[43] (VML). Obtaining access to the VML is relatively straightforward. An application is required, describing the purpose of the project and the credentials of the researchers involved, and specifying the datasets of interest. Allowances do need to be made for the time taken for the ONS to approve the project (in Ipsos MORI's experience this has ranged from two to six weeks, depending on the workload of the ONS).

[42] Total Factor Productivity (TFP) and average labour productivity (GVA per worker)
[43] The VML is a secure environment through which ONS allows researchers to access key administrative microdata for the purposes of research and evaluation.

Researchers may also be required to attend a training session on the disclosure rules associated with access to the data, which can also introduce a time lag.

For the ABS microdata to be exploited in an evaluation, it is necessary to identify the members of the treatment and comparison groups. As access to the data is strictly on an anonymised basis, it is not possible to achieve this directly within the VML. This can be resolved as ONS provide a data linking service through which they can link details of the individual firms of interest to the Inter-Departmental Business Register (IDBR) through the following process:

Figure 3.2: Data linking stages to match ESOS organisations to firm productivity & expenditure



Source: Ipsos MORI

Some further considerations for final additional stages to this approach are:

- **Agreement to data linking**: If there is interest in linking survey responses to the microdata (as there will be for the evaluation of ESOS), it is important to obtain the respondents' consent. The mainstage evaluation survey achieved a 62% consent rate for survey baseline data and responding company details to be linked to other datasets holding company-related information
- **Importing other data**: One potential hazard that requires consideration is whether the terms of licenses for other data permit their use in the VML. It is noted that BEIS has

permitted ND-NEED data to be used in the VML in the past which would be needed to compute energy efficiency metrics (energy consumed per unit of output).

- **Timelags**: There are some lags associated with the availability of ABS data. At the time of writing, ABS data was available to 2013, and data relating to 2016 is unlikely to be available through the VML until 2018. As such, primary research will likely be needed to provide nearer term measures of the effects of interest

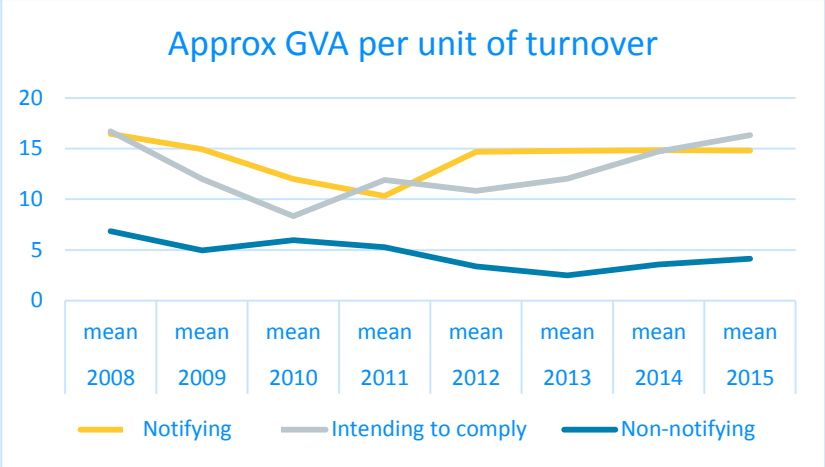
There are also some issues of coverage that need to be considered if this approach was also to be applied to assessing firm-level impacts for organisations employing lead assessors. The ABS only provides longitudinal data for firms with 250 employees or more which may be problematic for assessing any comparison organisations which are SMEs. The Business Structure Database (an annual snapshot of the IDBR that is used as a sampling frame for the ABS) could provide supplementary records of employment and turnover for SMEs (and financial sector firms which are also not represented) and could be accessed through the same process as the ABS. However, the data can be subject to lags of up to three years, making it challenging to accurately isolate the periods pre- and post-ESOS[44]'.

Findings from feasibility testing of this stage: The necessary permissions were obtained to access the VML and conduct testing on the data linking match rate to ESOS population data, as well as to run off example variables from the ABS to test the feasibility of generating baseline data in this way. The results of this testing are very positive: 96% of ESOS-obligated organisations have been matched into the ABS microdata. Longitudinal data was able to be run from the ABS, matched to ESOS-obligated records with a 'notifier' or 'suspected non-notifier' flag on variables such as, costs of energy in running of the business, capital expenditure and GVA per unit of turnover.

Key learning: The feasibility work has proven the ease of matching to VML datasets. While the ABS therefore offers a very strong dataset for many of the ESOS outcomes of interest, it is of more limited value should an impact evaluation strategy be chosen that focuses on SMEs (as a counterfactual) or which is not able to make use of the ND-NEED or raw meter data for energy consumption (the ABS only provides a proxy through its collection of energy spend data).

[44] The Business Structure Database is based on live data from VAT and PAYE returns and the results of Business Register Employment Survey. These values are not 'time-stamped, although we understand that ONS are changing their approach in this respect.

Figure 3.3: Example longitudinal variable from Annual Business Survey for ESOS population



Source: ONS[45]

3.5 Measurement of outcomes through primary surveys

As set out in Table 3.1, some of the key outcomes of interest will also be suited to collection through primary survey research with obligated organisations and assessors. Primary surveys could be used to provide data on energy efficiency investments and uptake of behavioural energy efficiency measures in particular.

3.5.1 Learning from an initial baseline ESOS evaluation survey

A telephone survey of 871 ESOS-obligated organisations was conducted as part of the wider evaluation contract delivered to BEIS by Ipsos MORI and UCL over 2015 and 2016[46]. This survey not only collected data for the process evaluation but also captured early indications of impact on key measures and set a baseline[47] that can be updated through a longitudinal survey (71% agreement was obtained in the baseline survey for company and respondent details to be stored for future research over the next 2 years).

[45] This statistical data from ONS is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

[46] Full report discussing initial ESOS evaluation survey findings available here: https://www.gov.uk/government/publications/energy-savings-opportunity-scheme-esos-evaluation-of-the-scheme

[47] The timing of the survey in August 2015 (driven by availability of Environment Agency data of ESOS obligated organisations) meant baseline measurements had to be captured retrospectively post the compliance deadline. The survey asked, for example, expenditure on energy efficiency measures in 2015, before collecting this information for 2016.

The key impact measures collected in the baseline survey included:

- Energy efficiency measures installed
- Capital spend on energy efficiency measures (£)
- Energy consumption (converted into kWh, but captured with in kWh, gWh or £s)
- Proportion of fleet that is electric or low carbon
- Number of employees involved in energy management or energy efficiency
- Level of priority placed on energy efficiency by organisation overall and at board-level (perception based)
- Behavioural action (e.g. energy use reduction goals, action plan in place, training)
- Certification to ISO 14001 or 5001

The proportion providing consent to re-contact for a follow-up survey and for data linking to administrative datasets providing further organisational information are presented in Table 3.3 below.

Table 3.3: Consent rates achieved during baseline ESOS evaluation survey by ESOS notification status

	Re-contact consent rate (sample size)	Data linking consent rate (sample size)
Notifier (inc. intend-to-comply)	71% (585)	61% (497)
Non-notifier	62% (31)	68% (34)

Source: Ipsos MORI

Effective primary research design to mitigate challenges in capturing ESOS outcomes: Challenges encountered in the design and administration of the baseline primary survey are discussed here to provide learning for the design of future surveys, or analysis of longitudinal measures.

Maximising participation: Obtaining participation from non-notifying organisations (either those who do not believe they qualify for ESOS or who are, as yet, non-compliant with the scheme), proved extremely challenging[48]. This is likely to affect the ability of a future impact evaluation to make robust comparisons between compliant and non-compliant groups on the basis of primary survey data, and emphasises the importance of making best use of secondary datasets (such as those already discussed in this section).

Table 3.4 below sets out the difference in response needed at the baseline survey between organisations who had and had not notified compliance (at a confidence interval of 95%):

Table 3.4: Response difference needed to identify significant differences between organisations' notification status in baseline ESOS evaluation survey (at 95% confidence level)

	Sample size	10% / 90%	30% / 70%	50% / 50%
Notifier (inc. intend-to-comply) vs. Suspected non-notifier	821 vs. 50	+/-8.3	+/-12.7	+/-13.8

Source: Ipsos MORI

Where outcomes data is most suited to primary data collection (for example, uptake of behavioural measures or perceived reputational outcomes), incentivising participation of non-notifying organisations may need to be considered to maximise response to any follow-up. Other best practice techniques should also be adopted, such as providing advance warning of the survey and providing a motivating introduction to the survey which explains the value of participating. The filtering of questions to smaller groups should be avoided as far as possible to prevent sample sizes at specific measures becoming smaller. While substantial analysis is still likely to be limited, some measures may be tracked over time; where a longitudinal sample is used, smaller sample sizes are required for the same statistical power compared with cross-sectional samples[49].

Maximising valid responses through flexible questions: piloting of the baseline survey found it to be important to offer flexible response options in order to maximise the

[48] Completed interviews were achieved with only 3% of 1,719 initial sample leads.
[49] Under a longitudinal design the relevant variance is within-subject (differences between repeated measures of the same organisations) rather than between-subjects (i.e. sampling error). The variance between the same units of interest at different points of time is generally considerably smaller than between samples of different units.

proportion of respondents providing a response on a given question. This was achieved, for example, by allowing energy consumption to be recorded either as units of energy demand (kWh or gWh) or as units of financial spend (in £s), and to be recorded for either 2015 or another recent accounting period (with the year end of this alternative accounting period collected). Similarly, offering banded ranges for capital spend on energy efficiency measures (rather than asking an open question for an exact amount) reduced non-response from 59% to 21%.

Maximising comparability of data capture: as the role and remit of respondents to the survey varied considerably across the sample (from facilities managers, to finance directors, and from those based at one operational site to those based in central shared services teams) collecting outcomes measures based on a comparable scale of organisational structure posed an analytical hazard. Following piloting, the baseline survey handled this by:

- Asking respondents about the scope of their responsibility for energy management (whether at a single site, some but not all sites or all sites).
- The size of the sites at which they have responsibility for energy management relative to the size of all sites across their organisation (asking this about energy spend, overall floor area, annual turnover and number of UK employees to account for respondent having varying levels of knowledge on some of these criteria and to maximise responses).
- Checking with respondents whether the response refers to all their organisation's sites or only those they are responsible for (e.g. for number of employees involved in energy efficiency, or fleet size).
- Conversion factors can then be applied based on these metrics to ensure key outcomes measures are based on comparable organisational scales (this was conducted successfully on energy consumption data for example during the analysis of the baseline survey).

Capturing investment (or planned investment) in energy efficiency measures: Requires significant prioritisation as a fully comprehensive list would be vast and highly varied according to relevance by sector, building and operations type and prior energy efficiency history. For the baseline survey, a prioritised list of 26 measures was agreed with BEIS using Salix Finance's comprehensive list of measures for which they hold persistence factors (a measure of the lifetime of each measure) and estimated savings in kWh/m2. Questions described above collecting floor area, and the scale of response relative to overall organisational scale, can be used to analyse these questions.

Maximising recall on key measures: Key measures (such as capital investment on energy efficiency, installed or planned measures, energy consumption or vehicle fuel efficiency) were specified to respondents in an advance data sheet (sent by email once first contact with a relevant organisational representative had been made by telephone[50]). While such an approach has been successful on previous Ipsos MORI business surveys (for example, collecting financial data), it was found to be of relatively limited value in the ESOS survey: many respondents wished to participate in the survey immediately (rather than review the datasheet and agree an appointed time for the survey), which may reflect many being in the role of Facilities or Buildings Manager, many of whom may not often be desk-based. This is likely to have affected data quality, with high proportions of non-response at some important outcome measures (25% for energy spend, 21% for capital expenditure on energy efficiency measures, for example).

Measuring impact across full in-scope population: The evaluation team faced challenges to identifying individual organisational entities that were in or out of scope of ESOS, which led to surveying of parent-level organisations only at the baseline evaluation stage. This limits the evaluation's understanding of impact issues across subsidiary-level organisations. This is particularly key to understand given the potential to notify at an aggregated level of company structure and some uncertainty about whether, and how, impacts anticipated for the scheme are realised among organisational levels less involved in the compliance process (for example, less close to the audit activity or recommendations report).

In light of the challenges discussed here, a combination of primary and secondary data collection on key measures is recommended (for example, complementing survey reported consumption and expenditure measures with ABS data, as discussed in section 3.1.5).

3.6 Further sources to explore ESOS outcomes

3.6.1 Considering value of pre-compliance tracker survey for ESOS impact evaluation

In the year leading up to the scheme compliance deadline, DECC ran a tracker survey with around 100 large organisations per wave over six separate waves. It has been examined to assess the extent to which it provides baseline measurements on outcomes of interest. Its value to the impact evaluation is limited, however, as it predominantly collected process-orientated data (capturing awareness and understanding of ESOS and tracking the progress of organisations towards compliance, including the completion of energy

[50] An alternative approach, previously used by Ipsos MORI, would be to send the datasheet alongside the advance email providing advance warning of the research. To avoid potentially reducing response rates by seeming overburdening, this approach was not taken for the ESOS survey.

audits). While additional questions to capture pre-observation measures on a number of key outcomes were added by Ipsos MORI to the final wave of this Tracker survey (December 2015), the small sample sizes involved in this survey mean it has limited value for tracking organisational priorities and actions over time[51].

3.6.2 Gathering outcome measures through qualitative methods

While many of ESOS' outcomes can be monitored through quantitative measures as set out in this section, there is also merit in considering qualitative approaches to understanding the realisation of some of the scheme's intended outcomes. Three key outcomes suited to qualitative data collection are:

- uptake of behavioural energy efficiency measures;
- perceptions of reputational impacts from ESOS compliance (or the associated uptake of energy efficiency measures and strategies – particularly in the case of ISO 50001 certifications);
- making estimations on potential effects of ESOS on suppliers of energy efficiency services and technologies through exploration of ESOS-obligated organisations' plans for future investment, auditing or energy consultancy.

During the parallel process evaluation of ESOS, which also explored initial indications of impact, these types of outcome were explored through qualitative in-depth interviews. A key learning from this stage which is recommended for a full impact evaluation is the value of exploring these outcomes qualitatively with multiple representatives within an organisation (through a case-study based approach). This facilitates engagement with employees with a range of roles and remits and with varying levels of engagement in compliance regulation and energy efficiency issues. This can add valuable further insight to primary data collected through surveys, which often are targeted at the lead contact involved in the scheme's compliance, or someone responsible for energy management issues. These respondents may not be able to offer a complete picture of how some impacts, particularly those based on behavioural change, are taking effect.

3.6.3 Review of available data on assessor registers for use in impact evaluation

Data enabling analysis of the supply of ESOS Lead Assessors has been accessed through 14 Lead Assessor registers, including details of 947 Lead Assessors. All registers provided the name of accredited assessors, however, the availability of other data fields was variable. Postcode data, (used to provide evidence for the geographical spread of assessors compared with qualifying undertakings) was available for half of the registers.

[51] Permission to re-contact organisations participating in the Tracker was collected but the available re-contact sample for whom pre-compliance outcome-related data was gathered is only 77 organisations (being based on just one wave in December 2015).

The registration date of assessors, (used to analyse the level of supply in the assessor market), was available for six of the 14 registers. Further information, such as the sectoral expertise or qualification held by assessors also varied in completeness from register-to-register.

3.7 Further considerations

3.7.1 Wider externalities

The impacts of ESOS on wider externalities are separated into impacts on greenhouse gases, air quality and health impacts. All indicators will be obtained from data on energy consumption divided by fuel from the ND-NEED database or the quantitative survey in the case of transport fuel consumption:

- **GHG emissions**: Impacts of ESOS on GHG emissions can easily be quantified by computing avoided GHG resulting from energy reductions (which are computed as part of the organisational level impacts) or their monetary value based on BEIS carbon values (noting there may be traded and non-traded savings).

- **Air quality**: Impacts of ESOS on air quality (e.g. SO2, NOX, and PM) can be computed at a basic aggregated and averaged level from energy reductions, by applying emission coefficients for the saving in specific fuels (as long as fuel information is able to be captured through survey questions about changes in heating systems, fleet characteristics etc.). This will not take account of spatial or temporal factors affecting air quality, however

- **Health impacts**: Similarly, sophisticated indicators for impacts of ESOS on health are complicated to compute due to health impacts being a function of local concentrations of pollutants as well as saved emissions. As such, it is suggested that the application of coefficients converting emissions into quality adjusted life years (QALYs) saved are likely to be insufficiently robust to merit inclusion in an ex-post cost-benefit analysis of the policy.

The feasibility of measuring ESOS' impact on energy market volatility has also been considered. By reducing peak-demand, ESOS may reduce the occurrence of scarcity events and therefore have a calming impact on price. Evaluating this impact with confidence is however problematic and for this reason it is not included within the suggested impact evaluation indicators. Insights on this matter could be obtained from the findings of the ongoing evaluation on Electricity Demand Reduction (EDR).

3.7.2 Control Variables

As discussed in Chapter 2, there is overlap in both the obligated (or targeted) population, as well as the outcomes of interest, between ESOS and a number of other non-domestic

energy policies (for example, CRC). It is important to capture exposure to these policy initiatives among ESOS-obligated organisations so that these can be treated as control variables in the analysis of outcomes and enable the effects of ESOS to be isolated as far as possible.

It is anticipated that exposure to other policy initiatives can obtained through the Environment Agency's records of ESOS-obligated organisations which also fall within the CRC or CCA populations (with the Environment Agency also the scheme administrator for these) as well as potentially other schemes. However, this information was not available to the evaluation team at the time of writing. As an indication, however, in the survey of ESOS-obligated parent organisations, 73% of respondents reported that their organisations participated in one of the following overlapping schemes or policies (CRC, EU ETS, CCA, Climate Change Levy, Mandatory Greenhouse Gas reporting, EPC, DECs).

Any BEIS administrative data for these other policies that can be made available to the full impact evaluation team could also be used to confirm the overlap between the ESOS population and eligible populations for other policies, or recipients of grants.

3.8 Summary of approach to measuring ESOS outcomes

The summary below provides a final review of the outcomes of ESOS that should be targeted for measurement through a full impact evaluation, based on findings from this feasibility study. The summary details the most appropriate source and method of measurement and provides an indication of the likely data coverage and quality of this outcome data for the ESOS-obligated population, reflecting the discussions around challenges and learning set out above. The summary focuses on the outcome measures set out as high priority in Table 2.1.

Table 3.3 Summary of ESOS outcome measures

Outcomes for obligated organisations	Most likely data source	Data coverage	Likely quality of measurement
Investment in energy efficiency measures (inc. fuel efficiency)	Annual Business Survey (supplemented by primary surveys)	**High for large firms** (but no coverage of SMEs obligated under ESOS)	**Medium**: ABS provides overall capital investment only, detail of energy or fuel efficiency measures from survey data likely to be incomplete
Implementation of behavioural efficiency measures	Primary surveys	**Medium**: lower for suspected non-notifiers & subsidiaries	**Medium**: self-reported data
Energy and fuel consumption	Annual Business Survey (pending further analysis of ND-NEED)	**High for proxy measure** but no direct data coverage	**Medium**: combined proxy measure based on spend
GHG emissions	Annual Business Survey (pending further analysis of ND-NEED)	**High for proxy measure** but no direct data coverage	**Low**: no direct observation, approximated from energy use
Productivity	Annual Business Survey	**High for large firms** (but no coverage of SMEs obligated under ESOS)	**High**
Abatement costs (wider environmental effect)	Annual Business Survey (pending further analysis of ND-NEED)	**High for proxy measure** but no direct data coverage	**Low**: no direct observation, approximated from energy use
Non-measurable outcomes			
Reputation	Perception only based measurement, not suitable for quantification		
Air quality and treatment costs for related morbidity	Cannot be directly observed, only approximated on basis of change in energy consumption but robust attribution to effects of ESOS is unrealistic		

4.0 Impact Evaluation Options

This section considers how, and through what method, any observed change in the outcome measures detailed in Chapter 3 can be attributed to ESOS, as opposed to another policy or driver. This includes a discussion of the range of impact evaluation options that might be applied to develop a robust assessment of the causal effects associated with ESOS on its key anticipated outcomes. In doing so, detailed consideration is given to the issues that might be involved in constructing an appropriate counterfactual group of non-ESOS organisations, including the potential to apply a range of econometric techniques to assess the impacts involved. A summary of the options appraised is presented in Table 4.3

4.1 Defining a Counterfactual

4.1.1 General Principles

A credible quantitative assessment of the impacts of ESOS will require the selection of an appropriate group of organisations that were not obligated under – or did not comply with – the ESOS regulations. Ideally, this comparison group should share both observable and unobservable characteristics with the group of organisations obligated under the regulations.

It may be possible to conceive alternative counterfactual scenarios for the impact evaluation - such as the implementation of the alternative policy options considered in the impact assessment. However, the adoption of such approaches is considered unrealistic as part of an empirical ex-post evaluation, since it would be infeasible to gather the evidence that would underpin the assessment as no organisations have been exposed to these alternative policy options.

4.1.2 Options for Constructing a Comparison Group

The primary effects of ESOS will be observed amongst organisations obligated under the regulation. These regulations are mandatory for large organisations and as such, do not create the typical challenges associated with self-selection into 'treatment' (i.e. the worry that ‘treated’ organisations differ from ‘untreated’ organisations in ways that are systematic and correlated with the outcomes of interest). However, the universality of the obligations creates a different set of challenges in that there are no large organisations that are not exposed to the intervention from which a counterfactual group can be formed. Given this constraint, the following possibilities have been explored:

a. Organisations not obligated under ESOS regulations

Comparisons between organisations obligated under ESOS and those that are not (but are otherwise equivalent in terms of their observable and non-observable characteristics) would be the optimal basis for a comparison group - supporting an intention-to-treat analyses[52]. Given the universality of the obligations amongst large organisations, it is not possible to define a counterfactual group of organisations that meets these criteria. However, there are various groups of organisations that are not obligated under ESOS regulations that could form the basis for a comparison group (though not necessarily equivalent to those obligated under ESOS).

i) SMEs

SMEs that are not within an enterprise group of a large firm are not obligated under ESOS and could be used to form a counterfactual group of firms to assess the causal effects of the regulations. The population of (or a sample of) SMEs can be identified from Companies House records[53] accessible through commercial data providers such as Experian, Bureau Van Dijk or Dunn & Bradstreet. Alternatively, this group could also be identified in the Inter-Departmental Business Register (IDBR) and matched to ND-NEED datasets).

Clearly, these firms can be expected to differ in systematic ways to large organisations and in ways that could be correlated with the outcomes. For example, SMEs are more likely to be exposed to the types of constraints in financial markets that may inhibit their ability to finance a programme of energy efficiency investment (which may lead comparisons between the two groups to overstate the impacts of ESOS). On the other hand, SMEs are less likely to have been exposed to other policy initiatives such as the EU-ETS, CCAs, and CRC. If these initiatives have been effective in exhausting the energy efficiency options available to those subject to these obligations, then comparisons could overstate the impact of ESOS.

These difficulties could be partly avoided by restricting analysis to SMEs that are obligated under ESOS by virtue of their membership of an enterprise group with at least one large organisation. There are 2,700 such organisations in the population of organisations

[52] ITT analyses are robust to issues driven by self-selection into treatment.
[53] Though details of employment and turnover will not be available for smaller firms that are only required to report their assets and liabilities. These details, however, can be uncovered by linking the relevant CRNs to the Business Structure Database via the ONS Virtual Microdata Laboratory.

obligated under ESOS. There is also a pool of 2.4m SMEs from which a suitable counterfactual group could be drawn[54].

Within the SME population there are also SMEs that have been subject to the CRC Energy Efficiency Scheme which could be included as a matching variable (when conducting Propensity Score Matching regressions). This would allow for some measures of non-transport energy use to become available through matching to the CRC scheme data, although this would be a small subset of the SME comparison group.

However, there may be residual concerns for an approach based on SMEs that are members of a wider enterprise group which arise from:

- **Comparability of the two groups**: For example, SMEs with large parents may face less acute constraints in financial markets if they are able to offer parent company guarantees. Such differences could not be addressed by sample selection alone.
- **Data availability**: As illustrated in Chapter 3, the availability of administrative data is less comprehensive than for large organisations.
- **Generalisability**: The focus on SMEs would reduce the generalisability of findings (accounting for 17 percent of the population obligated under ESOS).

ii) SMEs with a large foreign parent

Further refinements could potentially be found by focusing on a subsample of SMEs as a source of the comparison group. In particular, SMEs in an enterprise group with a large foreign parent (but no large operating unit in the UK) are not obligated under ESOS. These firms could be assumed to be broadly equivalent to SMEs with a large domestic parent (and it may be possible to assume two groups of firms share unobserved characteristics).

Exploratory research linking Environment Agency records to the Amadeus company hierarchy dataset was highly successful (as discussed in Chapter 3) but challenges have been uncovered in identifying these groups of firms specifically within this data. These may not be insurmountable as they relate to the way in which search queries can be run which may be overcome through a bespoke request[55].

[54] ONS Business Demography

[55] Amadeus is a powerful data source with extensive and mostly exhaustive data on registered companies, but its search patterns have been found to be limiting. The best search approach within the database is to provide a list of URNs and then choose which variables to extract for each URN. This was attempted manually, collecting all SMEs in the UK first and then feeding the list of their ultimate global parents back in, but the query was too large as Amadeus does not permit queries over 100,000 items.

Analysis on this basis is therefore feasible in principle but could carry a number of hazards:

As with the above, the results would not be generalisable across the ESOS population and could not provide the basis for a general cost-benefit analysis (i.e. external validity is low).

- It will be possible to control for some observed differences between organisations (e.g. sector, size, number of sites, local temperatures). However, unobserved differences between groups may persist (e.g. differences in managerial practices driven by cultural distance from the parent) and could be problematic if this is linked to energy efficiency behaviour. This difficulty is likely less acute than for more general comparisons between SMEs.

- SMEs are also less well represented in the longitudinal secondary datasets that might be exploited in an impact evaluation (as described above).

iii) Large organisations not obligated under ESOS

There are some large organisations that are not obligated under ESOS (e.g. universities that derive less than 50 percent of their income from private sources). These organisations systematically differ from obligated parties in relevant ways (e.g. they will predominantly be based in the education sector) and would not offer a good source for comparison group.

b. Non-compliant organisations

An alternative counterfactual strategy would be to compare compliant with non-compliant organisations, based on compliance status at a given historical point in time, noting that the Scheme Administrator is continuing work to bring obligated organisations into compliance. While this approach could offer more generalisable results, there are two main hazards that reduce its validity:

- **Selection bias**: In spite of its regulatory status and the risk of penalty for non-compliance, 100% compliance with ESOS has not yet been achieved. Those that comply are therefore likely to differ in systematic ways from those who do not. For example, those electing not to comply may give less attention or priority to energy efficiency issues than those that do, implying that they may have been less likely to deliver energy efficiency improvements even in the absence of ESOS (resulting in an overstatement of the effects of the regulation). This selection bias means there is a strong risk of distortion in comparisons between compliant and non-compliant organisations (to a more significant extent than may be encountered with selecting a comparison group outside the obligated population).

- **Assignment of treatment**: this counterfactual approach relies on being able to confidently assign any given organisation into either a treatment (compliant) or control (non-compliant) group. However, complexities in the information available that describes the compliance behaviour of organisations means there is a lack of clarity in this assignment. The rest of this section discusses the source of this challenge, and its implications for the impact evaluation, in further detail.

iv) International Comparators

- The final possibility considered is the selection of a sample of organisations in other international territories that were not exposed to the ESOS regulations. This would have attractions in that it may be possible to develop a sample of large organisations that can be considered equivalent to those obligated under ESOS. Despite these tractions, the use of such a counterfactual would create potentially intractable analytical hazards. For example, such a sample could not be developed from organisations located within other EU Member States (each of which will have implemented an equivalent policy). There are also substantial differences in the wider policy context between nations, alongside problems of simultaneity where changes in that context have occurred alongside the implementation of ESOS. As such, scope for international comparisons have not been explored further.

Challenges understanding compliance behaviour of in-scope organisations

Flexibility built into ESOS (aimed at minimising burden by accommodating different organisational circumstances), allows an Ultimate Parent Group to disaggregate for the purposes of notifying their compliance with ESOS. This section sets out the implications and challenges for using Ultimate Parent Group or organisation entities for constructing a counterfactual group.

By 5 December 2015 the Scheme Administrator, the Environment Agency, had received c. 4,000 compliance notifications, c,2,500 intent to comply late notifications and c.400 do not qualify notifications. The current estimate is that c. 6,300 Ultimate Parent Groups (including all their subsidiaries) are obliged to participate in ESOS. The number of notifications to date however is c. 6,870, i.e. more than the number of Ultimate Parent Groups. This means that more than one notification can be provided by each Ultimate Parent Group.

The Scheme Administrator continues to work to identify organisations that are required to participate in ESOS and have not so far submitted notifications. This has included writing to approximately 3,000 organisations which had not submitted a notification by the deadline or notified the Environment Agency of their intention to submit late.

Therefore, analysis of future notification data will depend on how these continuing steps by the scheme administrator to bring organisations into compliance result in further revisions to the notification data. The types of circumstances which may change the number of available organisational entities for analysis include whether:

a) The full extent of a notifying organisation's energy use and corporate structures have been taken into account in their notification and hence whether the full group is 'compliant'
b) The Environment Agency find some organisation's notifications are not valid (e.g. if based on inaccurate or invalid data)
c) Organisations currently self-declaring as 'Do not qualify' are later deemed eligible for the scheme by the scheme administrator
d) Organisations who have not yet notified compliance are or are not later confirmed to be 'Do not qualify'.

Furthermore, organisational structures are complex and this adds to the challenge of potentially constructing a counterfactual from organisational entities. Our initial analysis of the notification data (disaggregating the ESOS population and linking Company House Reference Numbers into the Amadeus dataset) suggests that there are 16,860 unique organisational entities[56] represented. However, this figure would be subject to additional analysis and confirmation once work by the scheme administrator was completed.

The greatest hazard posed by the challenges set out in this section is the uncertainty over the compliance status of any given organisational entity within an ESOS-eligible corporate group. Once the scheme compliance bodies complete work to bring non-compliant organisational entities into compliance, it will become clearer whether comparisons could be made between compliant organisations and any non-compliant organisations (if any remain).

4.1.3 Summary

As described above, the universal nature of the ESOS regulations creates a number of challenges for the selection of a comparison group to support robust causal analysis. Approaches focusing on a subset of SMEs obligated under ESOS will likely carry smaller risks in terms of distortions created by unobserved differences between groups, though with the trade-off that the results could not be generalised beyond the effects of ESOS on

[56] This is based on 13,914 unique Company House IDs identified in the Environment Agency's original ESOS population file, which increases to 16,860 when it is matched with the Company House IDs from the Environment Agency compliance notification file (dated 30th June 2016) - i.e. 2,946 additional organisations are identified as either notifying organisations or are listed within the notification forms as highest parent companies included within the notification

small organisations. In principle, comparisons could be made between compliant and non-compliant organisational entities to offer more generalisable findings regarding the impacts of ESOS, though this will remain an option only for the interim until non-compliant organisations are brought into compliance. Additionally, this would introduce more acute problems of selection bias. The following section explores options as to how these issues created by the choice of counterfactual could potentially be addressed through the application of appropriate analytical techniques.

4.2 Impact evaluation options

This section describes the options for econometric analysis that have been considered and provides an assessment of the extent to which they might be feasibly applied in an impact evaluation of ESOS. The robustness of these options is described in reference to the Scientific Maryland Scale (SMS); a five-point scale developed by Sherman et al (1997) to provide an indication of the methodological quality of impact evaluations and the reliability of estimates different approach produce of the causal effects of a policy under evaluation[57]. The numerical scale ranges from 1, for studies based on simple cross sectional correlations to 5 for randomised control trials.

4.2.1 Option A: Before and after (Interrupted Time Series) (Level II)

A before and after approach – or Interrupted Time Series (ITS) - is a straightforward method. It involves an impact evaluation based on the comparison of the value of one, or more, time series before-the-policy and after-the policy for organisations affected by the intervention. These types of method do not integrate a counterfactual group of organisations in their design and at maximum, might attain Level II on the Maryland Scale of Scientific Methods. As such, these methods only offer convincing evidence of impact if there is a reasonable degree of confidence that there are no external factors that might be thought to be influencing the outcomes observed. However, if a before and after design is not achievable, then more robust methods can be ruled out, so it is an important option to consider.

The feasibility of before and after methods are largely determined by the extent to which the outcomes of interest can be observed before and after the intervention. The analysis set out in Chapter 3 provides an optimistic view regarding the extent to which longitudinal data will be available for an impact evaluation (for example, energy consumption of

[57] See 'Review of Evaluation in Government,' Gibbons, McNally and Overman for National Audit Office, 2012 (page 12) for discussion of the Maryland Scale of Scientific Methods. Available at: https://www.nao.org.uk/wp-content/uploads/2013/12/LSE-Review-of-selection-of-evaluations-with-appendices1.pdf (accessed July 2016).

obligated or compliant organisations), and it is considered that a before and after approach is feasible in principle.

However, there are a range of factors that are likely to inhibit the extent to which this approach would yield robust findings. This is mainly due to challenges with the internal validity of this approach depends on there being no factors beyond the policy that influence the outcome of interest (this problem stems from the absence of a counterfactual). This assumption cannot apply in the context of ESOS (e.g. energy consumption will be determined by a wide range of other variables, including general economic growth, wider financial market constraints, broader energy efficiency policies, and external shocks):

- **Exposure to other policies**: A number of ESOS companies are affected by a number of other policies, e.g. CCAs, CRC and ETS but this should not be a problem as these companies are easily identifiable so that one can remove them from the sample (with the approach applied only to those companies affected only by ESOS). Changes in other policies influencing the operation of "ESOS-only organisations" have been limited in the period of interest. For example, we do not expect that energy performance of firms in the construction sector has been considerably influenced by the change in the zero-carbon commitments, as this policy is expected to impact the constructed buildings not the operation of the firms. The impact of the changes in the feed-in tariffs might be sizeable but an assumption is made that BEIS (via Ofgem) will be able to provide the location and size of all FiT and RHI installations so that this could be factored into the analysis.

- **Simultaneity with other policies**: Over the compliance period (and during the period in which compliant organisations may have been considering whether to pursue recommendations made to them as a result of ESOS), there have been changes to the energy efficiency policy landscape, however. This includes the announcement of reforms to improve the energy tax and reporting regime, including abolishing the CRC and absorbing the price signal into the Climate Change Levy[58]. It was also announced that a new simplified energy and carbon reporting framework will be introduced by April 2019. The timing of ESOS compliance alongside these shifts means any effects of ESOS will be inseparable from the effects of these changes. Additionally, ESOS compliance took place in the run up to the EU Referendum vote, which may also have affected energy efficiency investment decisions.

[58] HMRC, Climate Change Levy: main and reduced rates, 2016, https://www.gov.uk/government/publications/climate-change-levy-main-and-reduced-rates/climate-change-levy-main-and-reduced-rates.

- **Timescales to impact**: One set of concerns is related to the precision of the estimates from a before and after approach and the fact that wide standard errors may capture part of the impact of the policy. The use of ITS is also somewhat problematic in the case of interventions that are implemented slowly and take time to diffuse, two of the likely features of ESOS. It is also worth bearing in mind that the time delay associated with the implementation of the energy-saving measures after compliance is unknown. Hence, it is not immediately clear when a break in the series is likely to be observed due to ESOS. This is complicated by the fact that one would use annual data in the analysis.

In addition, challenges in defining the compliant and non-compliant population also create possible practical difficulties (e.g. tracking the energy efficiency behaviour of 'compliant' organisations could be challenging).

On balance, it is felt that concerns related to the use of a before and after strategy are significant and that without a strong case for internal validity this approach needs to be discarded as estimates of ESOS impacts obtained could differ in terms of magnitude and sign from the real effects of the policy. In light of the above, designs including the integration of a counterfactual will clearly be favoured.

4.2.2 Option B: Matching with Longitudinal Panel Methods (Level III)

An improvement on the first option would be to infer the causal effects of ESOS from comparisons between those that were obligated under (or complied under) ESOS regulations with those that were (did) not. This could be constructed using any of three viable counterfactual groups identified earlier in Chapter 4.

Such a strategy would raise concerns that the treatment and counterfactual groups of interest differed in systematic ways and in ways that are connected to the outcomes of interest. These could potentially be addressed through applying the following analytical techniques:

- **Matching methods**: Observable differences between groups (e.g. sector, characteristics of premises) could be largely accounted for through the application of statistical matching methods (such as Propensity Score Matching). These approaches involve an assumption that it is possible to identify all characteristics influencing assignment into treatment that may also relate to the outcomes of interest. A core challenge for the implementation of matching methods will be the weak availability of control variables beyond basic organisational characteristics such as sector, size, and past observations of the outcome variables (as set out in Table 4.1). While a rich set of controls have been collected for compliant organisations through the survey forming part of the process evaluation, only limited observations are available for a counterfactual group. In addition, there are also issues regarding unobserved differences between treatment and control that cannot be addressed through matching

methods. Such issues would be most acute in comparisons between compliant and non-compliant organisations (as explained in Section 4.1.2).

Table 4.1: **Review of potential matching variables**

Potential matching variable	Rationale	Potential data sources	Limitations
Sector	**Related to level of energy-intensity: process-evaluation found, for example, non-energy intensive sectors (such as retail), or office-based sites to be less engaged in energy issues and ESOS than those in manufacturing or with industrial sites**	**Companies House records**	
Premises characteristics (including floor area, building type/purpose)	**As above**	**Primary survey data would be needed for premises-level data**	**No large-scale survey of SMEs collecting this data, and response rate challenges limit data coverage for non-compliers**
Energy billing arrangements (i.e. whether combined with building rental/maintenance charges or separate)	**Related to engagement in energy use: process evaluation found**	**As above**	**As above**
Tenure of premises	**Can be connected with energy billing arrangement, as well as permissions to carry out building fabric improvements**	**Primary survey data would be needed**	
Prior installation of energy efficiency measures	**Related to energy use engagement but also potential to take further action**	**Primary survey data would be needed**	The baseline survey conducted during the parallel process evaluation highlighted data coverage and accuracy issues due to variable respondent knowledge
Participation in other energy efficiency policies (CRC, CCA, EU ETS, CCL Levy, Mandatory GHG reporting)	Related to engagement in energy use (from prior reporting requirements) and potential for further action	Environment Agency data, where scheme administrator of these other policies, and BEIS data	This data has not been assessed as it has not been available to the evaluation team for review
Involvement in public sector contract tendering	Related to reputational drivers for compliance and taking action	Publicly available frameworks	Identifying ESOS-obligated organisations may be challenging if a unique identifier is not included (e.g. the Companies House Reference Number)

Source: Ipsos MORI & UCL

- **Fixed effects**: The availability of longitudinal data would allow the application of panel methods designed to be robust to any unobserved but time invariant differences between SMEs and large organisations in the matched sample (i.e. fixed effects or difference-in-differences methods). This raises an additional practical issue that will be encountered in any evaluation design incorporating SMEs in the analysis, in that one of the key administrative datasets (the Annual Business Survey) only provides longitudinal data for large organisations.

In technical terms, this approach has reasonable internal validity (attaining up to Level III on the Maryland Scale) but low external validity (i.e. the results of the evaluation strategy could not be extended to the population of organisations obligated under ESOS, and could not be used to drive a cost-benefit analysis).

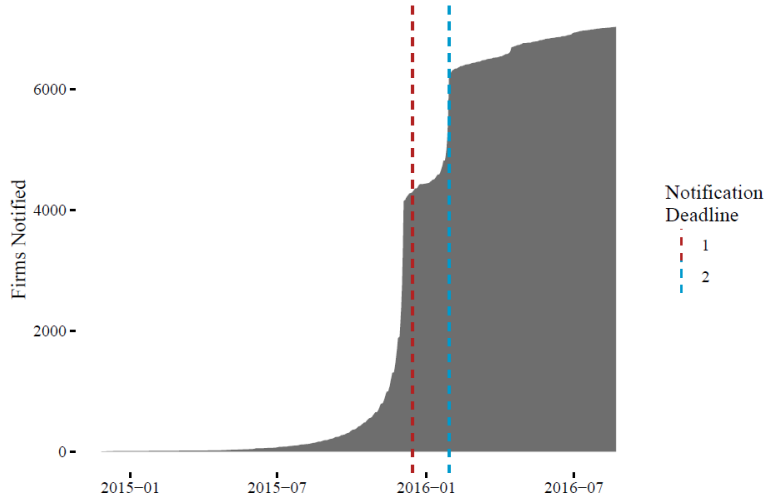
4.2.3 Option C: Pipeline Methods

This scoping study considered the viability of an evaluation strategy that could be feasible if there was staggering in the date at which obligated organisations notified their compliance – for example, if not all the organisations obligated under ESOS Phase 1 notified compliance in 2015/16, but instead some did so at later points between 2016/17 and 2019. This source of temporal variation could potentially be exploited to assess the impacts of the policy, with those organisations notifying their compliance later used as a counterfactual for those notifying in 2015/16 (and non-compliant organisations excluded from the analysis). In principle, such a strategy could address many of the issues associated with selection bias, on the basis that those notifying their compliance at later dates could be treated as equivalent to those notifying compliance at an earlier stage. Such a strategy also offers generalisable results regarding the effects of ESOS on the compliant population.

However, further scrutiny of the assumptions underlying this approach raise some concerns regarding its feasibility in the near term:

- **Degree of temporal staggering**: Analysis of compliance notification data has shown that there is not a high degree of temporal staggering in compliance. As shown in Figure 4.1, the staggering is short term in nature (months rather than years). 33% of all who provided notifications by July 2016, complied in the final week before the 5th December 2015 deadline, with a further 30% of notifications logged between 6th December and end of January 2016 in the grace period for compliance via audit activity. As most data will only be available on an annual basis, it may not be possible to leverage temporal variability to assess the impacts of ESOS. However, there may be scope to apply the method in a future evaluation of the programme (as organisations can comply with the next ESOS cycle at any point between 2017 and 2020), although in the first compliance phase, organisations also had the option of notifying before the deadline.

Figure 4.1: Date of ESOS compliance notifications



Source: Environment Agency compliance database, August 2016

- **Systematic differences between groups**: The validity of the approach depends on an assumption that there are no systematic differences between organisations complying at different points in time. This assumption may be hard to justify, since those organisations that did not comply in 2015/16 were in effect in breach of their legal obligations and are likely to differ in unobserved ways to those that met the compliance deadline. While the concerns could be somewhat minimised through using fixed effects models, this underlying issue is likely to prevent this approach attaining a level of robustness higher than Level III on the Maryland Scale. Again, such an assumption may be more straightforward to justify over the next ESOS cycle.

4.2.4 Option D: Instrumental Variables (Level IV)

As discussed in Section 4.2, there will likely be systematic differences between compliers and any non-compliers with the ESOS regulations. However, a robust comparison may still be feasible if it is possible to identify factors that influence the likelihood of compliance but have no direct causal link with the outcomes of interest. This strategy, known as an instrumental variables (IV) approach, can offer both unbiased estimates of the effects of interest (though only for those organisations whose behaviour has been influenced by the instrument). Further information on this approach is given in Annex 1.

Qualitative research feeding into the process evaluation suggested that ESOS assessors had a prominent role in determining awareness of the obligations. It was hypothesised that the strength of links between ESOS assessors and obligated parties would decline with spatial distance. If so, then those obligated parties in areas with higher densities of accredited assessors would be more likely to be aware of the obligations, and in turn, more likely to comply. As the location decisions of ESOS assessor are not determined by

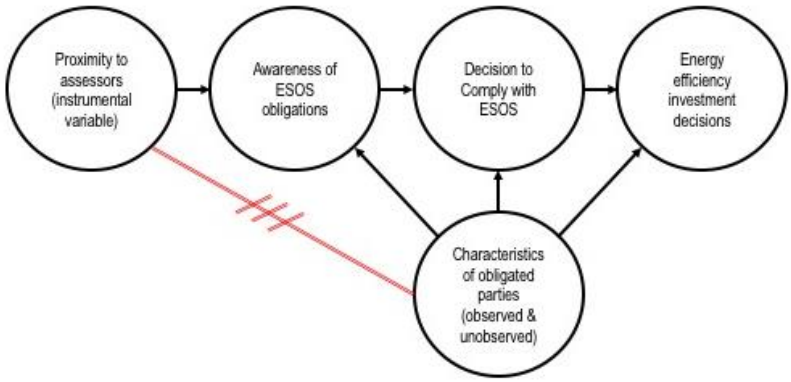
the characteristics of obligated parties, they would not be directly causally connected to the other factors that may influence energy efficiency behaviour (i.e. unobserved factors that would distort comparisons).

Implementation of this approach would be via Two-Stage Least Squares, involving the following:

- A regression of compliance status against density of assessors (using a probit or logit model). This would provide estimates of the likelihood of compliance status conditioned by the density of local assessors for each organisation in the sample. These estimates, by assumption, would be uncorrelated with other unobserved factors driving compliance behaviour that also drive energy, efficiency behaviour.
- The estimated likelihood of compliance derived from the above regression in each organisation in the sample replaces the variable describing compliance that would be included in an Ordinary Least Squares regression model. If the key assumptions hold, this 'instrument' would be uncorrelated with confounders and produce unbiased estimates of the impacts involved.

Therefore, it was hypothesised that local density of ESOS assessors could potentially be used as an instrumental variable to provide an unbiased estimate of the causal effects involved. This hypothesis is shown in the figure below:

Figure 4.2: Proximity to Assessors as an Instrumental Variable



Source: Ipsos MORI & UCL

For an instrumental variable approach to be valid, two key assumptions need to be satisfied. Firstly, the instrument needs to be correlated with the likelihood that organisations comply with the ESOS obligations (i.e. it needs to be a 'strong' instrument). Secondly, the instrument needs to be valid – i.e. uncorrelated with the unobservable characteristics of organisations that may determine both their compliance status and their energy efficiency investments.

Some tests of these assumptions have been conducted by combining data on compliance with data on the location of ESOS assessors (obtained from the registers held by the 15 accreditation bodies) which is visualised in the figures below. Figure 4.3 shows all registered assessors. Each point is a registered assessor, and the points are semi-transparent so where the colour is darker, there is a higher density of assessors. Figure 4.4 shows ESOS-compliant organisations and non-notifying organisations. Compliant organisations are shown in green and non-notifiers are shown in red. Where each colour is darker, there is a higher density of corresponding terms[59].

Figure 4.3: Registered assessors

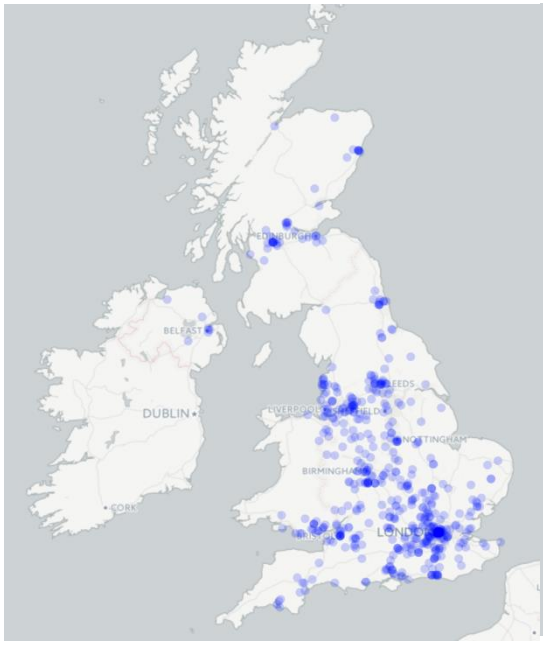


Figure 4.4: ESOS-obligated organisations

Source: Ipsos MORI & UCL

Tests were run to establish whether there was a significant difference in the mean number of registered assessors in a 10km range between compliers and non-notifiers[60]. Using this e definition, this test rejected the null hypothesis that the true difference in means is equal to zero[61]. This provides initial evidence that there is a statistically significant relationship between density of proximal assessors and propensity to comply, implying that instrumental variables could be a feasible option.

[59] If the required data is able to be made available to the evaluation team, it is recommended that further mapping is conducted which maps the geographical distance between each obligated organisation and the Lead Assessor they commissioned to support their ESOS compliance activity.

[60] Compliers being those who had stated compliance; 'do not qualify'; or intention to comply in their notification.

[61] Based on Welch Two Sample T-Tests, generated a p-value of 0.028.

The validity of the approach could potentially be contested. There is a question regarding the direction of causality in this instance: it may be possible that assessors chose to locate – or become accredited - in certain areas because there were a high proportion of organisations that became aware of the obligations through other means (although qualitative evidence collected in the parallel process evaluation does not support this hypothesis – suggesting instead that assessors were a key, and often the first, point of information about the scheme). In this event, the instrument would be correlated with the unobserved characteristics of organisations that determined compliance. It is not possible to test for the direction of causality in this instance.

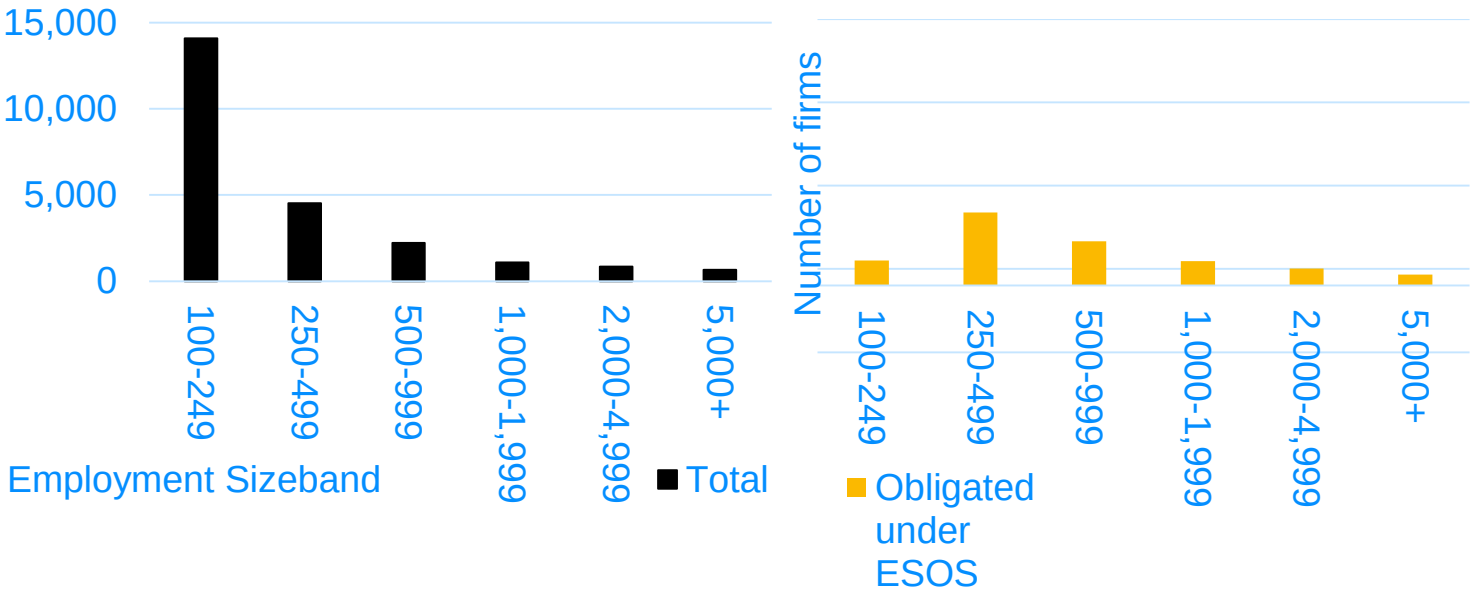
A full assessment of the validity of the instruments suggested can be done only when implementing the instrumental variable estimation. This is also subject to the challenges discussed in section 4.1.2 on constructing a counterfactual group and the work by the scheme administrator to bring organisations into compliance.

4.2.5 Option E: Regression Discontinuity Design (Level IV)

The design of the policy creates a discontinuity at the employment and turnover thresholds above which organisations become obligated under ESOS regulations. On the assumption that differences between organisations in the vicinity of those thresholds can be considered as random (i.e. organisations with 249 employees can be broadly considered equivalent to organisations with 251 employees), comparisons between those just above and below the eligibility threshold could potentially yield an assessment of the causal effects involved that approximate the robustness of a Randomised Control Trial (provided certain assumptions are satisfied). This strategy is known as Regression Discontinuity Design.

Some organisations with fewer than 250 employees may be obligated under ESOS either because their turnover exceeds the minimum eligibility threshold or if they are a member of an enterprise group where another undertaking meets the eligibility criteria. There is not a ‘sharp’ discontinuity between those in and out of scope of the ESOS regulations (and it is estimated that around 11 percent of SMEs with 100 to 249 employees were obligated under ESOS). This is not problematic, though it would require ‘fuzzy’ methods that require a discontinuous increase in the likelihood an organisation is obligated under ESOS at the minimum threshold. This requirement is demonstrated in Figure 4.5 below.

Figure 4.5: Total Firms and ESOS-obligated Population, 2015*



Source: Ipsos MORI.

* This figures combines UK business counts produced by the Office for National Statistics with Environment Agency records of the ESOS-obligated population from March 2016. The latter are drawn from Companies House filings and may include defunct organisations (and may partly explain why the number of obligated firms exceeds the population at some sizebands). Non-qualifiers may also be excluded from the EA source data.

However, applying a fuzzy RDD approach to ESOS may violate some of the key assumptions that determine its validity:

- The statistical precision of RDD strategies can be partly contingent on the density of observations at the relevant threshold. Analysis of Environment Agency records suggests that there are around 250 organisations obligated under ESOS with 250 to 260 employees. The low density of observations at the threshold could prove problematic in analysis[62].
- The RDD approach assumes that no other relevant factors change suddenly (or discontinuously) at the relevant thresholds of interest. This assumption is violated in the

[62] See for example, Statistical Power for Regression Discontinuity Designs in Education, Deke and Dragoset June 2012, which suggests that a sample size of 3,500 would be needed to reach the same level of statistical precision as a Randomised Control Trial RCT with a sample size of 250, owing to statistical uncertainty associated with the relationship between the assignment variable and treatment status, sample loss associated with selecting an optimal bandwidth for analysis, and adjustments to standard errors to account for random misspecification error.

case of ESOS as, despite the existence of a sharp and clear-cut threshold for inclusion in the policy, the 250-employees qualifying criterion coincides with the demarcation between Small and Medium Enterprises and (SMEs) and large organisations. There are a range of other regulations that change at the threshold between SMEs and large organisations. One example is the more relaxed treatment of SMEs and large organisations under the EU State Aid regulations as well as under other UK schemes such as R&D credits or policies to stimulate lending to businesses. These policies have the potential to lower the external costs of funds for SMEs, and may make lower payback investments more attractive than for large organisations.

- The RDD approach also assumes those receiving the treatment cannot manipulate their treatment status. Organisational employee size (and therefore their eligibility under ESOS) is within the control of the management of obligated organisations however. Realistically, organisations are unlikely to determine headcount based on ESOS eligibility but they are much more likely to do so because of the different regulation regimes on offer for SMEs and large organisations. Unfortunately, results from the RDD model are biased regardless of the motivating reason leading to the manipulation of the forcing variable (which is headcount in the case of ESOS).

- As noted, data availability will be weaker for SMEs (i.e. those with fewer than 250 employees) creating a further practical constraint on the suitability of this method. Finally, while internal validity of RDD methods is high, the results could not be generalised beyond those organisations with 250 employees (i.e. because the findings cannot be generalised beyond the groups just above and below the threshold).

4.2.6 Further considerations: **Different pathways to ESOS compliance**

An alternative approach that could be applied within the population of ESOS compliant organisations would produce counterfactuals based on the different ways in which compliance can be achieved. This avoids the need to tackle problems related to systematic differences between compliant and non-compliant organisations or source data for organisations not covered by the policy.

Analysis of the Environment Agency's published notification data (from August 2016) has been conducted to consider whether there are systematic differences between organisations choosing a certain compliance pathway rather than another. This analysis shows that ISO 50001 certification was consistent across organisations of different sizes, although firms in the construction sector were more likely to comply via the ISO 50001 route than others. While these are factors that can be assessed empirically and controlled for, there are likely to also be significant unobserved differences between these groups. Survey evidence from the parallel process evaluation found organisations giving higher priority to energy efficiency were more likely to comply via the ISO certification, and that

this was also related to perceptions of cost and time burden and the likely reputational value to organisations.

As 95% of compliant organisations commissioned external auditing services, the analysis will also be restricted by the small sample of ISO 50001 certifiers.

The key drawback to this approach, however, is that it can only measure the additional impacts of a number of pathways to compliance, after taking one as the control group or baseline, rather than the impact of ESOS as a whole (but it may provide valuable information to adjust the policy and improve its effectiveness).

4.2.7 Alternative non-statistical approaches (Theory-based methods)

Given some of the challenges identified above in relation to a quasi-experimental impact evaluation of ESOS, it is also worth considering the extent to which a non-statistical approach could be employed to answer the key impact evaluation questions set out in Chapter 1.

This section discusses what could be offered by one such method adopting an overall framework based on realist principles. Within this, it considers the value of using a contribution analysis approach to assess the strength of the 'contribution story' linking ESOS's inputs and activities to outcomes and impacts at a specific organisational level, and it also discusses the potential value of Qualitative Comparative Analysis (QCA) to analyse across organisational cases. The latter could, in particular, be valuable to make sense of the analysing organisational case-study findings given a large volume of qualitative evidence which would be central to a non-statistical contribution analysis approach to the impact evaluation while offering a means of generating generalisable conclusions about the policy.

Value offered by a non-statistical approach: Adopting realist principles[63] in the impact evaluation of ESOS would centre it around developing an understanding of what elements of the ESOS policy do and do not work in achieving its intended objectives, for whom and under what circumstances. Given the broad range of organisational types obligated under ESOS, this explicit consideration of context could be of significant value to the unpicking and understanding of impacts and the mechanisms through which they come about.

Given the challenges set out earlier in this section in determining specific counterfactuals, a contribution analysis approach would provide an alternative way to explore the impacts of ESOS and infer causality at a case-level. The focus would be on the extent to which observed results (likely to be gathered through survey data and in-depth organisational

[63] Pawson, R. and Tilley, N., Realistic Evaluation, Sage, 1997

case-study visits) are due to ESOS, rather than other factors, helping to disentangle the potentially complex set of drivers of the anticipated outcomes. This contribution analysis would centre on developing a suite of 'competing performance stories' that offer alternative explanations and routes to the outcomes anticipated or observed. Extraneous effects, outside of the scheme, that would need to be considered when exploring cause and effect relationships may include the variety of political, economic, social and technical elements discussed in Chapter 2.

Finally, conducting Qualitative Comparative Analysis (QCA) across organisational cases provides a structure for generating learnings at a broader level than each individual case assessing the wider implications. Case-study approaches alone can have weaknesses as while they provide great depth in a particular outcome in a particular context, it is difficult to interpret how generalisable their learnings are to other projects or contexts. In addition, if not carried out systematically and transparently, findings from case studies may not have external validity (i.e. be replicable by external researchers). Building realist principles into the QCA will ensure it focuses on unpicking success factors (either external or internal) that influence the realisation of outcomes.

Key additional steps in process: Adopting a realist framework to the impact evaluation would require further expansion of the Theory of Change development set out in Chapter 2, including the development of a set of context mechanism outcome hypotheses for testing. These CMOs provide the basis for the competing performance stories (explanations of the alternative ways that the outcomes of interest could come about – and guarding against a confirmatory bias towards the Theory of Change) that would be explored as part of a contribution analysis approach to understanding the role of ESOS specifically in leading to observed outcomes.

Data collection would likely be through longitudinal case-studies, following the progress of both compliant and non-compliant organisations over the four-year ESOS cycle (further discussion of a case-study approach, and learning from the employment of this method during the parallel process evaluation is presented in Chapter 3). Further qualitative evidence gathered from stakeholders (from within the assessor market as well as potential ESOS Lead Assessor candidates who chose not to be accredited under the scheme; from the wider energy efficiency products and services market; and those involved in other energy efficiency policies) would also be needed to explore a full range of perspectives on the achievements of the scheme and perceptions of its relative contribution against other driving factors (other energy and non-energy policies affecting large organisations, energy prices, the availability of energy efficiency advice outside of the scheme, and so on).

Qualitative comparative analysis (QCA) would then be considered as an approach to synthesising the qualitative findings across the case studies.

Risks and limitations to approach: The main disadvantage to taking a non-statistical contribution analysis based approach to an impact evaluation of ESOS is that it cannot robustly demonstrate a causal link or provide a quantification of the scale of the effects involved and will not produce results that can feed into a Cost Benefit Analysis.

4.3. Summary of considerations and recommendations

The quantitative evaluation options identified above have been scored (on a scale of 1 to 5) against several key criteria:

- **Potential robustness** (using the Maryland Scale of Scientific Methods as a guide)**Analytical confidence** (whether there are issues that may compromise the robustness of the approach)
- Level of **confidence in implementation**
- **Generalisability of findings** to the ESOS population
- **Data availability**

The relative weighting given to these criteria will be critical in selecting an impact evaluation method that best meets the needs for the evaluation, the type of evidence that is required and the key questions it seeks to answer (an emphasis on accountability over learning, or vice versa, could affect the direction of the evaluation, for example).

The analysis (presented in Table 4.2 and based on equal weighting of the appraisal criteria) suggests that there is no dominant evaluation option that is preferred over the others, owing to the variety of trade-offs that need to be made. More robust approaches tend to focus on SMEs and, therefore, are weakened by the level of data availability and lower levels of generalisability. Approaches seeking to obtain more generalisable findings or exploit greater levels of data availability for large organisations tend to carry analytical hazards or implementation risks (partly stemming from the lack of additional data in the monitoring information describing compliant and non-compliant organisations).

Based on the analysis presented in Table 4.2, and with each of the appraisal criteria given equal weight, a mixed evaluation strategy may be preferred combining the highest scoring options:

Option B: Matching and Longitudinal Panel Analysis using non-obligated SMEs as a counterfactual (including SMEs with a large foreign parent); and,

Option D: Instrumental Variables using proximity to lead assessors as an instrumental variable.

While the application of these approaches may offer BEIS some insight into the effects of ESOS, it will not be possible to offer a quantitative evaluation design that simultaneously meets all evaluation needs and minimises all implementation risks. It should also be noted that the implementation of these options will be to a greater or lesser degree dependent on administrative data that is only available with a lag (it will not be possible to collect longitudinal observations through survey research for a comparison group). This will have timescale implications for a future quantitative study: it may not be possible to determine the effect of ESOS in 2017 until 2018/19.

To help compensate for some of these issues, it is advised that non-statistical approaches are integrated into the design of the evaluation grounded in realist evaluation approaches. It may also be worth considering how far the implementation of Option A could add value to such a programme of research (which could involve collecting longitudinal follow-up observations from those surveyed as part of the process evaluation).

Table 4.2: Summary of appraised quantitative impact evaluation methods

Evaluation Option	Counter-factual	Potential Robustness	Analytical Confidence	Implementation Confidence	Generalisability	Data Availability	Notional Score
A: Before and After	None	Maryland Level II (2)	**Very Low (1)** Simultaneity with other policies and events highly problematic	**High (4)**	**High (4)**	**High (4)**	**15**
B: Matching with Longitudinal Panel Analysis	Non-Obligated SMEs	Maryland Level III (3) Intention-to-treat analysis	**High (4)** Low risk of selection bias driven by unobserved differences between groups	**High (4)**	**Low (2)** Only captures effects on SMEs	**Low (2)** No measures available from ABS for SMEs	**15**
	Non-Obligated SMEs with large foreign parents	Maryland Level III (3) Intention-to-treat analysis	**Very High (5)** Lowest risk of selection bias driven by unobserved differences between groups	**Medium (3)** Possible hazard regarding available sample sizes	**Low (2)** Only captures effects on SMEs	**Low (2)** No measures available from ABS for SMEs	**16**
	Compliers vs Non-Compliers	Maryland Level III (3)	**Very Low (1)** Likely severe issues regarding	Low (2) Substantial challenges in identifying individual entities	**High (4)** Captures effects on compliant organisations	**High (4)**	**14**

			unobserved differences between groups	for analysis of compliant and non-compliant units *			
C: Pipeline Design	None – later compliers vs early compliers	Maryland Level IV (4) Assumption that late and early compliers are equivalent	**Very Low (1)** High likelihood late compliers differ in systematic ways to late compliers	**Very Low (1)** Insufficient staggering of compliance to make this a viable option	**High (4)** Captures effects on compliant organisations	**High (4)**	**Not feasible**
D: Instrumental Variables	Compliers vs Non-Compliers	Maryland Level IV (4)	**Medium (3)** Residual issues regarding causality in the candidate instrument – possibility that location choices made in response to organisational awareness of ESOS	**Low (2)** Substantial challenges in identifying individual entities for analysis of compliant and non-compliant units *	**Medium (3)** Captures effects on compliant organisations that were influenced to comply due to their proximity to assessors	**High (4)**	**16**
E: Regression Discontinuity Design	SMEs	Maryland Level IV (4)	**Very Low (1)** Discontinuous changes in other policies at the ESOS eligibility thresholds	**High (4)**	**Very Low (1)** Captures effects on organisations at the eligibility threshold	**Low (2)** No measures available from ABS for SMEs	**12**

* Note also that the EA will continue to bring organisations into compliance, so there may be limited opportunities to carry out analysis on non-compliance organisations before they are brought into compliance

5.0 Economic evaluation

This section sets out a framework for an economic evaluation of ESOS, itemising the range of costs and benefits that could be included and highlighting the assumptions that may need to be applied.

5.1 Aims for a longer-term economic evaluation of ESOS

An economic evaluation of ESOS would take two key forms:

Cost-effectiveness analysis, exploring the unit cost of the impacts or results delivered relative to other similar initiatives,

Cost-benefit analysis, placing a monetary value on the impacts delivered and relates these to the costs involved.

5.1.1 Principles for economic evaluation

It is envisaged that an economic evaluation would be delivered as a full social welfare analysis in line with the principles of the HM Treasury Green Book, drawing largely on estimates of the causal effects of ESOS estimated through the impact evaluation:

- **Counterfactual**: The cost-benefit analysis would be designed to assess the incremental costs and benefits resulting from ESOS (i.e. those costs and benefits that would not have accrued in its absence). The reference case for the assessment would be a counterfactual scenario of no ESOS regulation. This is in alignment to the approach to the ex-ante appraisal of costs and benefits set out in the DECC Impact Assessment for ESOS as well as broader evaluation guidelines specified in the HM Treasury Magenta Book. This implies that the cost-benefit analysis arguably should be driven – as far as practicable – by direct estimates of the causal effects of the regulation derived through the parallel impact evaluation.

- **Costs**: All incremental opportunity costs should be included within a cost-benefit analysis. This should include costs incurred by both the public and private sector as a direct and indirect consequence of the regulation.

- **Benefits**: Equally, the cost-benefit analysis should include consideration of all improvements (and losses) in social welfare (which should be examined and

monetised as far as practicably possible). This implies a central focus on the final impacts of ESOS (as defined in the logic chain). While there may be important intermediate indicators of the success of the policy that will be addressed in the impact evaluation (e.g. behavioural change), these do not necessarily imply an improvement in social welfare unless they can be demonstrated to lead onto improvements in overall economic efficiency - including reductions in negative externalities (i.e. CO2 emissions). An assessment of benefits should also consider the extent of any opportunity costs (in the form of alternative investments that may have been crowded out by ESOS).

- **Uncertainty and persistence**: There will inevitably areas of residual uncertainty regarding the net benefits of the ESOS following any impact evaluation taking place in the near term. These uncertainties will largely relate to the extent to which any effects of ESOS persist, grow or decay into the future. Sensitivity analysis regarding plausible scenarios regarding the future will be required to estimate the present value of the cost and benefit streams involved.

An economic evaluation may also require the inclusion of a cost-effectiveness study depending on how far it is possible to quantify the downstream impacts of any energy efficiency investment or changes in behaviour induced by ESOS. For example, if it is not possible to fully monetise the social welfare benefits and dis-benefits of ESOS, it may be preferable to undertake a cost-effectiveness analysis relating the overall resource cost of ESOS to its causal effects in reducing energy consumption (£s per KWh saved, for example).

5.2 Costs involved in ESOS

5.2.1 Scheme management costs

The resources expended in the management of the scheme should be included as a cost within a cost-benefit analysis of the regulation. This assessment should span the costs incurred by:

- BEIS in the design, oversight and management of ESOS; and,
- Environment Agency in their role as scheme administrator (including costs of designing the accreditation regime and implementing audit activity – BEIS involvement in this process also to be included within the BEIS portion of scheme management costs).

Such costs might reasonably be assumed to be wholly attributable to ESOS and can be obtained directly from the Environment Agency and BEIS. There may be challenges in isolating the staff costs and overheads associated with the delivery of the scheme if these costs are not allocated to cost centres.

5.2.2 Costs of maintaining registers of assessors

Organisations responsible for maintaining public registers of accredited lead assessors are external to the Government and its agencies, and it may be more difficult to obtain this type of information as part of an economic evaluation. It is anticipated that primary research (in-depth interviews, and a data collection exercise) with the 15 organisations concerned would be needed to obtain these costs.

5.2.3 Accreditation costs

Suppliers of ESOS audits will incur costs in acquiring accreditation under the scheme. These costs will largely relate to the costs of any training received – including any opportunity costs associated with attending the training. It may be possible to measure this directly through research with assessors. However, while such costs might again be reasonably assumed to be wholly attributable to ESOS, there is some risk of double counting if the fees paid to organisations maintaining public registers cover both training and scheme management costs.

Assessors wishing to gain ESOS accreditation may also be required to pay more general fees to organisations maintaining public registers, for example to gain membership status. While costs associated with training should be included in the CBA (noting the challenges raised above), those associated with membership will not. The latter is also not a cost that is wholly attributable to ESOS as by becoming a member of one of the trade organisations it might offer other benefits too. This raises some complexities and so it is suggested that membership fees are excluded for a number of reasons:

- There will be double counting to the extent these fees pay for the marketing activities of the trade associations with respect to ESOS; and,

- Spill-over benefits to the assessor organisations as a consequence of membership are not being included as part of the CBA of ESOS, and including the broader costs involved in producing these benefits would have the potential to distort estimates of the present value of the policy.

There is also a question of how significant these costs will be. Evidence from the parallel process evaluation suggests that around 900 assessors obtained accreditation

at price between £100 and £1,000. This implies total accreditation costs of around £90,000 to £900,000 (trivial in comparison to a total compliance cost of around £100m), and the exclusion of these costs is unlikely to materially alter the conclusion of a CBA.

5.2.4 Marketing and promotion

It is anticipated that the suppliers of ESOS will incur costs in marketing their services to large firms and other organisations. While these costs may be partly attributable to ESOS, they could also be accompanied by reduced spending in other areas or displace other types of marketing activity (i.e. only additional resources invested in marketing should be included within the CBA). ESOS may also improve the efficiency of marketing activity if suppliers use ESOS to cross-sell other services. However, there is a question of how far such costs are likely to be sufficiently significant to merit the detailed research that might support their quantification, and it has been agreed with BEIS that consideration of these costs will be excluded from a later economic evaluation. Marketing and promotion costs incurred by BEIS and the Environment Agency will be included within the scheme management costs set out above.

5.2.5 Costs of compliance (internal and external)

The cost of compliance will comprise two main components:

Internal costs: Firstly, there will be internal costs for obligated organisations that will primarily take the form of the staff time absorbed in deciding on the chosen route to compliance, procuring and managing an ESOS audit, and/or managing the process of achieving the relevant ISO accreditation[64] as well as assisting any quality assurance audit activity.

External costs: Organisations are likely to incur costs in payments to external suppliers of ESOS audits or in ISO accreditation fees (assuming they did not choose to comply by accrediting internal workers as lead assessors).

These costs can be estimated within a Standard Cost Model framework. A SCM framework was therefore specified to capture these internal and external costs through the baseline survey conducted in August 2016.

There is a possible issue in that not all expenditures incurred to comply with ESOS will be additional. To achieve compliance, organisations may have reduced spending on

[64] Such costs could spill-over to non-obligated firms if they invest any resources in determining whether they are required to comply with ESOS.

other services (including energy consulting), reduced the size of internal energy management teams, or delayed small scale energy efficiency (or other) investments to procure the energy audits required for compliance. Those choosing to comply through the ISO route may have sought to do so in the absence of ESOS (though possibly at a later date). These effects will reduce the net expenditures incurred by obligated parties because of ESOS, but could be expected to have opportunity costs in the form of the lost productivity associated with these expenditures.

Given the size of the organisations obligated under ESOS, it is anticipated that expenditures on audits will be trivial relative to their overall consumption of goods and services. It is anticipated that it will not be possible through econometric analysis to identify the causal effect of ESOS on the relevant measures that might feed into a CBA. In light of this, a conservative approach has been agreed with BEIS in which gross expenditures on ESOS - as estimated through the survey - are treated as wholly additional. Estimates of these costs are provided in the parallel process evaluation.

5.2.6 Costs of responding to ESOS-led recommendations (voluntary)

As identified in the evaluation framework, ESOS is likely to produce two types of downstream effect on organisational behaviour:

- **Investment in energy efficiency projects**: Organisations could incur capital expenditures to implement energy efficiency projects identified in the audit. To the extent that they have not been funded through reductions in spending in other areas, the present value of those expenditures should be included in CBA. It is anticipated it will be possible to estimate the effects of ESOS on both energy efficiency investment (from observations gathered through surveys) and overall fixed capital formation (through the Annual Business Survey). To estimate the present value of those expenditures, assumptions will need to be made regarding the rate of capital depreciation. An exercise in researching the lifespan of different types of energy efficiency technologies (including low emissions vehicles) will be required to refine these assumptions as part of developing this framework. Additional investment may also be accompanied by two types of other cost that would in principle require inclusion in a CBA:

- **Hassle costs**: For significant programmes of capital works, obligated parties may experience disruption that reduces their efficiency over the period of their completion.

This would be experienced as a temporary fall in Total Factor Productivity (TFP) which can be estimated using longitudinal data from the Annual Business Survey[65]. It cannot be assumed that such hassle costs would be sufficiently significant to be visible in any econometric analysis (and in any case would be offset by any positive productivity gains achieved by the firms as a consequence of ESOS). Supporting evidence from the qualitative research and potentially the survey of obligated parties will be required to qualify any results.

- **Maintenance**: Investment in new technology may also be accompanied by additional on-going maintenance costs. These costs may not be additional (or indeed may involve lower maintenance costs) if the new technologies replace older capital items and it is considered that an estimation of the net effects of ESOS on maintenance expenditures may be unrealistic - would also be offset by positive productivity gains achieved by firms (though again, they could be explored through qualitative research).

- **Behavioural Change**: Alternatively, ESOS audits may identify opportunities for organisations to reduce their energy consumption through adjusting the behaviour of workers and the way they interact with existing building and transport technology. These might be billed as 'no cost' recommendations, though to the degree that the desired changes in behaviour absorb the time of the workers concerned, there could be losses in the form of reduced productivity. As discussed in Chapter 3, while effects in this area could be explored through qualitative and quantitative research, consideration will need to be given as to how these costs might be realistically separated from any positive productivity gains achieved as part of the impact evaluation (though it should be noted that the cost of behavioural change initiatives will be implicit in the estimates of the overall impact of ESOS on firm productivity as described below).

5.2.7 Service Innovation

The evaluation framework identified the possibility that those organisations supplying ESOS audits - particularly those delivering large volumes - could use the data gathered to develop new or improved service offerings. In principle, such investments in R&D would need to be included in a CBA. However, these effects are tangential to the

[65] See 'The Causal Effects of an Industrial Policy,' Criscuolo, Martin, Overman and Van Reenan, 2012 for an outline of how TFP estimates can be generated from ABS data using a perpetual inventory method to estimate the value of capital stocks.

underlying objectives of ESOS and the associated costs will be challenging to quantify. It is suggested that qualitative research with lead assessors is used to explore these types of cost and qualify the central results of the CBA.

5.3 Benefits of ESOS

5.3.1 Productivity gains

Any reductions in energy demand driven by investments in energy efficiency or behavioural change will result in an improvement in productivity through reducing the operational costs of the firm. Productivity improvements may also arise indirectly through improvements in employee productivity driven by enhanced lighting or building facilities. As noted above, aspects including maintenance costs and the costs associated with behavioural change programmes could have positive or negative effects on productivity.

The present value of these resource savings (or efficiency losses) should be treated as a benefit (or dis-benefits) in a cost-benefit analysis[66]. Several complexities require consideration:

- **Monetising cost savings**: Recent supplementary Green Book guidance developed by BEIS recommends monetising productivity gains driven by reductions in energy demand at the long-run variable cost of production, rather than the retail price (to exclude taxation and other transfers). The guidance supplies benchmark values that can be applied to changes in energy demand or consumption.

- **Crowding-out**: Depending on the rationality of decision-making processes, energy efficiency investments may ‘crowd out’ competing programmes of capital or R&D investment with the potential to raise productivity. This may be problematic if there is a narrow focus on energy demand or efficiency, as this would overstate the net effect of ESOS at the level of the firm.

- **Displacement**: If energy efficiency investments reduce the operating costs of obligated firms, they may be able to reduce their prices. In turn, this may allow them to expand their market share, potentially resulting in offsetting displacement of sales and profits from competing firms based in the UK. These effects cannot be – or are

[66] It is equivalent to an increase in aggregate supply.

highly challenging to - observe or estimate directly as it is close to impossible to identify the firms likely to be subject to the spill-overs of interest. A conservative approach is advised in which the primary focus is on the change in output driven by improved efficiency (i.e. excluding profits on greater production volumes[67]). This will understate the impact of ESOS where obligated firms are primarily competing with overseas producers (e.g. as is the case in automotive or aerospace manufacturing).

Productivity gains can be estimated in four different ways:

- **Indirect estimation**: It may be possible to estimate the productivity gains involved indirectly on the basis of the effect of ESOS on the type and scale of energy efficiency investments and behaviour change made by firms. This would link evidence on the nature of the investments made to secondary evidence on (1) the expected reductions in energy consumption associated with the technology or investments involved and (2) data on energy prices to estimate the cost reductions involved. This approach may offer a route to obtaining estimates of the likely impact of ESOS in the nearer term as it could be driven largely by survey evidence (and relevant baseline data has been gathered as part of the process evaluation) Complexities are also introduced by evidence that energy efficiency measures do not always yield the performance improvements they are technically capable of[68].

- **Energy intensity**: Alternatively – or additionally - estimates might be driven by direct observations of energy efficiency (energy consumed per unit of GVA produced). These would be derived by combining microdata in the Annual Business Survey and the ND-NEED datasets, and while providing superior evidence to the approach set out above, would only be available at a lag and the timing of any later econometric analysis will need to be considered as part of specifying the main-stage economic evaluation programme. Again, there would be a need to use data on energy prices to monetise these cost savings.

- **Expenditure on energy**: Estimates of cost savings could also be directly driven by observations of energy expenditures per unit of GVA produced (exclusively using Annual Business Survey results). These estimates will be priced at the retail price

[67] The overall change in output (GVA) can be decomposed into a efficiency and an expansion component as follows (where y = output, and n = employment): $\Delta y = \Delta \frac{y}{n} \cdot n + \Delta n \cdot \frac{y}{n}$.

[68] See 'Building Performance Evaluation Programme: Early Findings from Non-Domestic Buildings,' Innovate UK, November 2014.

rather than the long run variable costs involved. However, as highlighted in Chapter 3, it may be possible to treat the marginal percentage change in energy expenditure as close to equivalent to the marginal change in energy demand).

- **Direct observations of productivity**: Finally, ABS data could also be used to determine how far ESOS has led to changes in productivity at a firm level. While such results would be robust to any crowding out of competing investments, there would inevitably be a risk that the effects of ESOS would be too small to be visible in these firm level measures. Issues regarding the potential statistical power of any analysis will also require consideration in the development of this framework.

- **Qualitative research**: Qualitative research with applicants could also offer guidance on the nature of any productivity gains arising from ESOS (though clearly, such results could not be used within a cost-benefit analysis of the policy).

5.3.2 Turnover effects driven by Reputation

In line with HMT Green Book guidelines, it is recommended that any sales growth driven by reputational factors, that is not accompanied by an increase in efficiency, is excluded from the CBA (i.e. it is considered as simply displacement)[69].

5.3.3 GHG Emissions Reductions

Estimates of the value of abatement costs avoided can potentially be driven by estimates of the impacts of ESOS on energy demand. The required analysis would be largely assumption driven, with reductions in electricity and gas consumption converted into associated reductions in GHG emissions. These reductions in emissions would then be monetised using BEIS's Carbon Valuation framework. The assumptions used within these estimate calculations follow the Supplementary guidance to the HM Treasury Green Book on Appraisal and Evaluation in Central Government.

5.3.4 Air Quality

Air quality benefits are typically monetised in terms of their health benefits and will largely be driven by improvements in fuel efficiency in transport. However, the scale of

[69] The exception to this would be if consumers are willing to pay more for goods or services they believe have been produced in an environmentally friendly manner. This can create a social welfare improvement in the form of higher 'consumer surplus' (the value that consumers gain from trade: i.e. most consumers are willing to pay more than the market price for products and services, so they gain a 'surplus' which is the difference between what they were willing to pay and what they actually paid), although it is very challenging to robustly monetise such an effect.

health benefits will in turn be contingent on the spatial concentration of these fuel efficiencies (i.e. there may be limited health benefits if the fuel efficiency gains are made by drivers predominantly using rural roads). Without observations on the specific patterns of road use by organisations obligated under ESOS, it may be challenging to credibly monetise any air quality benefits.

5.3.5 Non-quantifiable gains

There are a range of benefits that may be challenging to quantify:

- **Energy security**: Reduced energy consumption by obligated organisations will place less stress on the electricity network, potentially reducing the extent of price rises during scarcity events.
- **Service innovation**: It will be highly challenging to measure and monetise any secondary effects of ESOS in terms of supporting service innovation in energy consulting and the development of novel energy efficiency saving measures more generally.
- **Supply chain multipliers**: A conservative assumption will be adopted that any supply chain effects created by investment induced by ESOS will be offset by general equilibrium effects (i.e. any increase in demand will place pressure on factor prices, reducing demand elsewhere in the economy).

5.4 Summary

The table below provides a summary of the costs and benefits that would need to be covered in a CBA of ESOS and the sources of evidence that would be required. Note that this analysis is contingent on the availability of robust estimates of the costs and benefits of energy efficiency investment. Given the uncertainties as to how far these estimates can be feasibly produced (see Chapter 3), there are questions regarding the feasibility of an economic evaluation of ESOS.

Costs / Benefit	Included?	Source
Costs		
Scheme management costs	Yes	Environment Agency and BEIS
Maintaining registers of assessors	Yes	Primary research with 15 accredited organisations
Accreditation costs	No	
Costs of compliance	Yes	Process evaluation
Cost of responding to ESOS	Yes	Impact evaluation (see Sections 3/4)
Marketing and Promotion	No	
Service Innovation R&D	No	
Benefits		
Productivity gains	Yes	Impact evaluation (see Sections 3/4)
Turnover effects	No	
GHG Emissions Reductions	Yes	Derived from estimated impact on energy consumption and BEIS Carbon Valuation Methodology
Air Quality	No	

6.0 Implementation strategy

This section proposes an implementation strategy for an impact evaluation of ESOS, based on the recommended approach presented in this report. It details the audiences and methods that would need to be employed across a preparatory and main stage phase, specifying the evaluation activities involved including timelines and risks.

6.1 Preparatory work

Prior to the mainstage impact evaluation of ESOS, it is advised that the following preparatory work is conducted. The following steps will help minimise some of the risks highlighted in the earlier sections of the report (as well as confirm feasibility), and ensure the impact evaluation is based on as complete and accurate scheme and secondary data as possible. Firstly, the scheme data will need to be updated to provide:

- **An updated understanding of the ESOS-obligated population**: the impact evaluation team will require an updated database of the population that were obligated under ESOS in 2015/16. It is anticipated that this would reflect checks by the scheme administrator into 'Do Not Qualify' notifications and into non-notifying organisations originally considered in-scope on the basis of publicly available company records.
- **An updated understanding of the compliance status of this obligated population:** an updated compliance notification database will be required following the EA's process of bringing organisations into compliance, so that it is clearer whether analytical comparisons could be made between compliant organisations/ subsidiaries and non-compliant organisations/ subsidiaries (if any are still non-compliant at that time).

Secondly, the following preparatory steps should be conducted on secondary datasets:

- **Identification of site-level addresses for ESOS-obligated organisations.**
- **Identifying SMEs in-scope of ESOS with large foreign parent companies**: to pursue this counterfactual option this will require further investigation into the

functionality of the AMADEUS dataset to identify these firms. The critical issue is determining whether (and the costs of) a bulk download can be provided.

- **Review of ND-NEED**: the current coverage of meter profile data in ND-NEED is a significant constraint on the ability of the evaluation to track longitudinal energy use for ESOS-obligated organisations. However, it is understood that there is work ongoing to review how the matching of meters to organisational sites can be improved to increase this level of coverage. It is therefore recommended that the team delivering the ESOS impact evaluation hold further discussion with BEIS around the use of this dataset and any further improvements that have been made or strategies that can be implemented to overcome this challenge.

6.2 Mainstage evaluation specification

This section sets out the key audiences to be engaged in an impact evaluation of ESOS, the recommended data collection approaches and how the evidence will be analysed and synthesised to answer the overarching evaluation questions.

6.2.1 Audiences

- **Obligated organisations**, including compliers and non-compliers.
- **Policy stakeholders**: including BEIS and DfT policy teams, Treasury stakeholders.
- **Scheme management**: including the Environment Agency, and the devolved regulators (Scottish Environment Protection Agency, Northern Ireland Environment Agency, Natural Resources Wales).
- **Assessor market**: including ESOS Lead Assessors, representatives of the organisations hosting approved registers, energy consultants not becoming accredited under the scheme.
- **External stakeholders**: energy efficiency technology manufacturers, installers.

6.2.2 Data collection approaches

Primary research strands: The mainstage phase of an impact evaluation of ESOS will involve the following key phases of primary research.

The data collection strands presented in Table 6.1 will provide data that would feed into an impact evaluation based on a before-after design with no quantitative counterfactual (achieving a robustness of Level 2 on the Scientific Maryland Scale). The inclusion of:

Table 6.1 Primary research specification for ESOS impact evaluation

	Rationale / Key evidence needs	Audience/specification	Recommended timing
Longitudinal survey	Capture outcome data from complier organisations: - Investment in energy efficiency measures - Implementation of behavioural measures - Perception of reputational outcomes - Linking any action taken to ESOS audit recommendations and expected savings - Understand what other factors have contributed to actions taken (assisting a more realist approach to analysis)	25- minute survey of **complier organisations**	Baseline survey from Autumn 2016 to be **repeated in Autumn 2018**: this is up to three years' post-compliance giving time for outcomes to have been realised, but prior to when compliance activity for the next ESOS cycle is likely to start (deadline December 2019)
Longitudinal case-studies	Understand how organisations react to ESOS recommendations and outcomes are realised (or hindered) through organisational decision-making & gather qualitative and perception-based outcome measures Key element of data collection for realist impact evaluation approach	**Compliant and non-compliant organisations**; expanding this strand to engage: - 10 x organisations from original baseline survey (providing long-term follow-up; 8 compliers, 2 non-compliers) - 10 x organisations from follow-up survey (to be reflective of range of outcome circumstances identified in this survey; 8 compliers; 2 non-compliers) To include qualitative consultations with organisational employees, external assessors, and related document review	**Autumn 2018** (as above)

Stakeholder in-depth interviews	To provide wide-ranging perspectives under realist approach requiring creation and testing of competing performance stories	Stakeholder groups set out above in section 5.1	**Initial engagement in late 2017** to develop further Theory of Change In-**depth interviews in late 2018** (providing an opportunity to test and reflect on survey findings, considering potential role played by ESOS vs other factors)

Source: Ipsos MORI

survey questions seeking attribution of ESOS to any organisational investments or changes made, could help overcome the lack of counterfactual to an extent, but the reliability of these estimates would be weak given they would be based on self-reporting.

Secondary source data linking: Two main strands of data linking activity will also form part of the mainstage specification. The steps involved in the implementation of these data linking activities are set out below:

Data linking to energy use data (Figure 6.1) – pending feasibility

Data linking to firm-level productivity and expenditure data (Figure 6.2)

Figure 6.1 Data linking stages to match ESOS organisations to energy use data

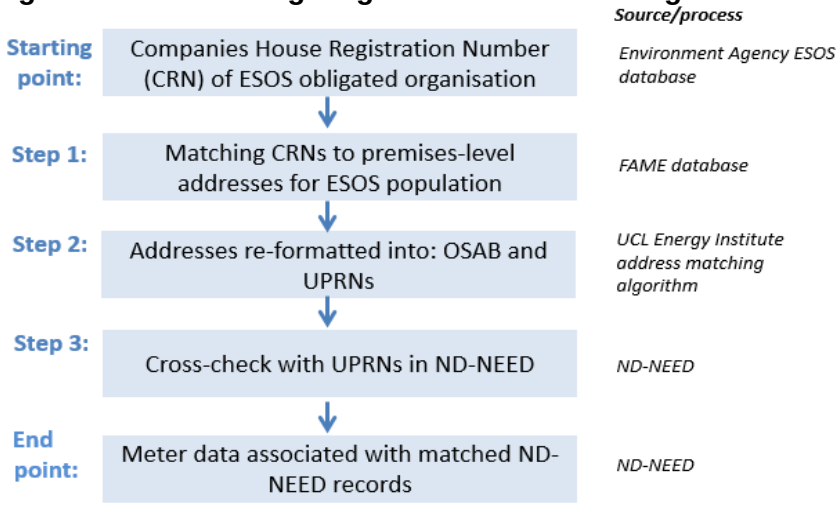
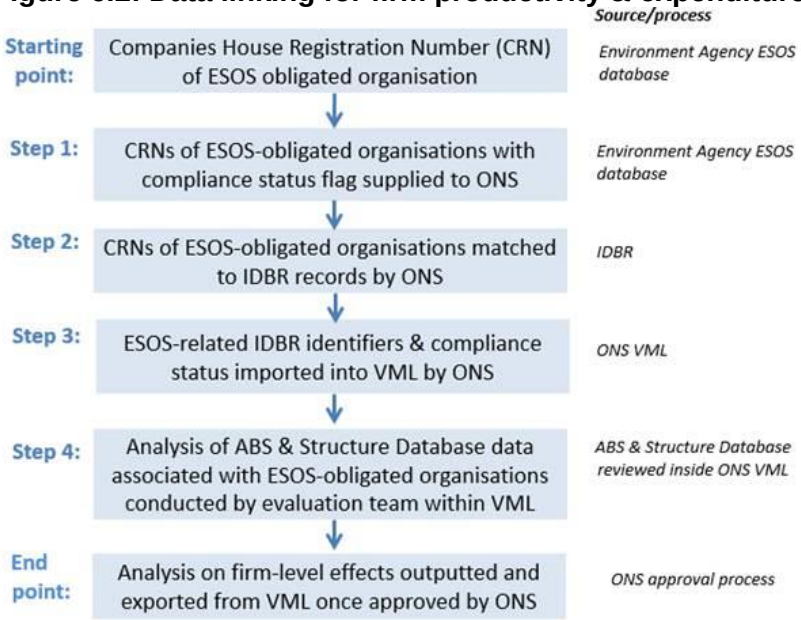


Figure 6.2: Data linking for firm productivity & expenditure outcome data



Source: Ipsos MORI & UCL

The aim of this data linking activity will be to track ESOS-related outcomes from a baseline measure from 2014/2015 to repeated measures in 2017-2019; this follow-up time period is

recommended as the parallel process evaluation has shown that compliant organisations were only in the initial stages of considering their ESOS recommendations by late 2016, and so it is unlikely that the effects of ESOS could become observable in data such as energy consumption and firm productivity until 2017 at the earliest.

The timing of implementing the data linking will need to take account of any lags in the publication of data. Based on the evaluation team's current understanding of the involved datasets, it is anticipated that the data linking can be implemented to the timescales set out in Table 5.2. This may mean a separate report based on this data is required following reporting of the primary evidence described above (which can be completed in 2018):

Table 6.2 Data linking implementation timetable

Data linking aim	Source	Data availability	Implication for timing of implementation
Data linking to energy use data	ND-NEED	Current available data for 2012 (published in 2015)	Baseline data for 2014/5 likely to be available in 2017/8 Outcomes data for 2017-2018: likely to be available in 2020-2021 (unless earlier data release is able to be negotiated with BEIS)
Data linking to firm-level productivity and expenditure data	ABS	Current available data for 2014	Baseline data for 2014/5 likely to be available through the VML in 2017/8 Outcomes data for 2017-2018: likely to be available in 2019-2020

Source: Ipsos MORI

6.2.3 Approach to analysis

Quantitative data analysis: Three main approaches to quantitative analysis will be applied to quantitative measures generated through the complier survey, as well as through secondary data linking and scheme monitoring information:

- **Raw data synthesis**: During the data management stage, raw data in SPSS and Excel format will be reviewed and sorted and frequency tables created where necessary.
- **Descriptive and inferential analysis**: Once data has been synthesised, descriptive and inferential analysis will be used governed by the theory of change. Monitoring information will be primarily descriptive providing summaries about scheme compliance rates, for example.
- **Time-series analysis**: In order to assess changes in perceptions and reported actions since the baseline period, time-series analysis will be employed.

Econometric analysis: Following completion of the secondary data linking stages above, econometric analysis can be implemented in line with the recommendations set out in section 4.4. This set out a mixed evaluation strategy combining:

- Matching and Longitudinal Panel Analysis using non-obligated SMEs as a counterfactual (including SMEs with a large foreign parent) - Option B; and,
- Instrumental Variables using proximity to lead assessors as an instrumental variable – Option D.

The anticipated timing of this stage will therefore be in 2020 (based on Table 5.2 above).

Realist evaluation analysis: In addition, Chapter 4 recommended the integration of non-statistical approaches grounded in realist principles to further aid an understanding of ESOS' impacts. The primary evidence feeding into such an approach would be case-study data which may be suited to a qualitative comparative analysis (QCA) approach (as described in section 4.2.7).

The principle behind QCA is to compare a number of factors across a number of cases in order to create a set of generalisable and transferable insights. This involves testing the theory of change by identifying the factors that are necessary to obtain the specified outcome or output within each context, in this case, within each case study. The extent to which the case study provides evidence that it has been able to achieve this outcome is then assessed by defining a set of scoring criteria and developing a dataset.

Table 6.3 exemplifies how we would expect this dataset to be defined when assessing a specific outcome.

Table 6.3 Example application of QCA to analysis of ESOS

Case	Factor A:	Factor B:	Factor C:	Outcome
	Level of priority placed on energy efficiency by board	Role & level of influence of ESOS lead within organisation	Satisfaction with quality of assessor & trust in ESOS recommendations	Implementation of behavioural measures
Organisation 1	0.66	0	0.66	0.66
Organisation 2	0	0.33	0.66	0.66
Organisation 3	0.33	0.66	0	0
Organisation 4	0.66	0	0.66	0.66

* 0 = no or weak evidence to support; 0.33 = some evidence to support; 0.66 = strong evidence to support

Source: Ipsos MORI

6.3 Synthesis of evaluation evidence

As demonstrated throughout this report, an impact evaluation of ESOS will involve the collection and analysis of a broad range of data collated through a number of different work streams. The analysis of evaluation findings will be guided by the systematic theoretical framework created through the theory of change (set out in Chapter 2). This defines an implicit set of hypotheses around the ways in which outcomes from the scheme may be realised that will need to be tested as part of the evaluation.

The quantitative and qualitative evidence (including both primary and secondary) gathered through the evaluation will need to be triangulated against this theoretical framework to provide a judgement as to how far the relevant hypotheses are supported (or not supported) by the results. In general, it is anticipated that the quantitative results will demonstrate how far the anticipated outcomes have been delivered, with the qualitative findings used to provide the evidence required to explain how a particular result has been achieved.

As outlined in Chapter 4 a variety of analytical approaches will be employed and a key consideration for the synthesis of the outputs from these multiple analytical strands is dealing with any contradiction in the findings produced. A ‘weight of evidence’ based approach is recommended which acknowledges that the evidence will likely vary in both its strength and variability, and will require an explicit appraisal of how strong an assessment might be made (driven by the consistency of findings across observed units and evidence sources, as well as the overall number of relevant observations).

6.4 Risks to implementation

This section sets out the risks facing the successful design and delivery of an impact evaluation of ESOS and considers how any risks may be mitigated or minimised to effectively answer the Key Evaluation Questions (KEQs).

6.4.1 Data Quality

The central risk to the feasibility of a quantitative impact evaluation of any description is the quality of the available data describing the population of organisations obligated under ESOS and records of those entities that were covered (and not covered) by compliance notices. Without accurate information, a set of assumptions regarding the coverage of compliance notices will need to be applied in an evaluation which could threaten both the quality of findings and the likelihood that quantitative approaches will detect the causal effects of interest. Improvements to the quality of the available information should be a priority over the interim period, as suggested in Chapter 3.

6.4.2 Linking to meter records

Reliable information on the energy consumption of organisations in both the treatment and any comparison group will need to be derived from the administrative data. This study has not yet shown that the challenges in linking records of organisational addresses to ND-NEED can be adequately overcome, though some improvement is possible by linking these addresses to raw energy meter records. An alternative is possible if the evaluation relies on measures of spending on energy taken from the Annual Business Survey (which will limit the focus to large firms and reduce the number of impact evaluation options available).

6.4.3 Overburdening businesses

In line with cross-Government ambition to reduce the administrative burden placed on businesses, it is important to minimise this also in the design and delivery of evaluation and research. Overburdening businesses may mean they refuse to participate in evaluation activity, reducing the quality and breadth of evidence collected in order to answer the KEQs. This would also have the potential to add to any negative perceptions of the scheme. Efforts should be made throughout the evaluation process to reduce the requirements placed on businesses. Examples of this include convenient scheduling of short telephone interviews, rather than requiring lengthy online, self-completion forms to be conducted. Secondary data and information collected by the scheme administrator should be maximised as far as possible – reducing wherever possible the extent of questions that would need to be asked directly to organisations. While the direct views of organisations remain a vital input into the evaluation, the evaluation team, in collaboration with BEIS, should continue to seek out further opportunities to minimise burden where opportunities arise (for example through combining primary research activity with other commissioned studies if timings and aligned objectives allow).

6.4.4 Lack of access to audit reports

Ideally the impact evaluation would have access to the ESOS audit reports in order to analyse the range of recommendations made and assess these against any evidence gathered around energy efficiency measures implemented since compliance with the policy. Given the audit process, and its outputs, is central in affecting whether or not organisations complying with ESOS pursue energy efficiency measures and behavioural changes this would naturally be a key piece of analysis informing the evaluation. In the absence of audit reports being centrally requested and stored (given the policy design) this evaluation will seek to collect key pieces of information about the audit reports during the primary research (for example, asking surveyed organisations to recall how many separate recommendations were made, what the overall expected costs and savings from these recommendations was estimated to be). Where possible, full analysis of audit reports will be conducted on those provided by organisations willing to share them with the evaluation team (for example, those participating in the case-study visits).

6.4.5 Quality assurance

An internal peer review process within the evaluation team is recommended so that members less involved in the design or administration of data collection provide a critical challenge function to the analytical outputs of the strand to help ensure the findings are defensible and stand up to scrutiny. It is anticipated that an external peer review process will be established by BEIS to provide a further level of independent quality assurance of the methods employed, analysis conducted and conclusions drawn.

Annex 1

Further Note on Instrumental Variables

What is instrumental variables?

Instrumental variables (IV) is a method for estimating causal relationships when a controlled experiment is not available. Specifically, IV can be used when there is selection on unobservables - when there are unobserved factors that may affect assignment of a treatment. If subjects can choose whether or not to partake, for example, and we cannot observe all factors that may influence this decision, then instrumental variables may be a good choice.

In more formal terms, if we have a linear regression model $y_i = \mathbf{x}_i'\beta + \alpha D_i + u_i$ where:

y_i is the outcome variable for subject i;

$\mathbf{x}'_i\beta$ is the matrix of observable covariates and corresponding parameter estimates;

D_i is a dummy variable, set at 1 or 0 if the subject has received the treatment or not (respectively);

α is the effect of the treatment on the outcome variable; and,

u_i is the error term.

Then, in an observational context, it is possible that D and u are correlated. Because unobserved variables affecting y will end up being subsumed into u, it is possible that there is a set of variables A that is unobserved and affects both y and D. If this is the case, then D will be correlated with u, which violates the Gauss-Markov theorem and leads to biased and inconsistent results from ordinary least squares (OLS).

In order to estimate a consistent and unbiased estimate of the effect of D on y, we use instrumental variables. Intuitively, the process involves finding some variable z which only affects y through D. That means $Cov[z, v] = Cov[u, z] = Cov[\mathbf{x}, u] = 0$ and $Cov[D, z] \neq 0$, or that y depends on z only through D and that D depends on z in a nontrivial fashion. This is usually estimated using two stage least squares (2SLS), which means estimating the following linear equations:

$$\hat{D}_i = \boldsymbol{\lambda}_i'\gamma + \rho z_i + \epsilon_i$$

$$y_i = \mathbf{x}_i'\beta + \alpha\,\hat{D}_i + u_i$$

where $\hat{D}_i$ represents the estimated value of D as predicted by some linear equation including z.

Department for Business, Energy & Industrial Strategy
3 Whitehall Place, London SW1A 2AW
www.gov.uk/beis