Annex to Rule Review Framework

Our approach to impact evaluations

January 2024

1 Introduction

1. In 2018 we published our framework for ex post impact evaluation (impact evaluation). We said that we would keep it up to date as our approach evolved. The Rule Review Framework (the Framework) sets out our approach to monitoring and reviewing our rules, including impact evaluations. We incorporated elements of our existing impact evaluation framework into the Framework. This Annex provides more detail about our approach to impact evaluation because:

    - impact evaluations are our most technical and most data-intensive type of reviews
    - we can use impact evaluations for interventions other than rulemaking
    - we want to retain and update some of the case studies and examples we set out in our 2018 impact evaluation framework as they flesh out our approach to impact evaluation

2. In this Annex we cover:

    - why we undertake impact evaluations
    - how we conduct impact evaluations
    - case studies
    - the key challenges

2 Why we undertake impact evaluation

1. We have committed to undertake at least 1 impact evaluation per year.

2. Impact evaluation:

    - Helps us getting a better understanding of what works or doesn't work, and why. The lessons learned through impact evaluation help to improve our policy making and inform future decisions.

    - Provides a strong **quantitative evidence base** for ensuring market changing regulations are proportionate, for repealing ineffective regulation or clarifying how our actions have added public value.

    - Helps to **improve the assumptions we make in our cost benefit analyses** (CBAs) by allowing us to compare post-intervention effects with those estimated before the event. We can identify areas where we systematically and materially under or overestimate outcomes of an intervention, and inform key assumptions in our CBAs.

    - Helps us to **understand and demonstrate our net impact** and the costs for firms. Those costs need to be set against the benefits we bring to consumers and society through our operations and policies.

3 How we conduct impact evaluations

1. The aim of our impact evaluations is to assess the impact of our interventions in the market. In particular, we look at whether we achieved the intended positive difference and whether there have been unintended consequences. We do this by focusing on the fundamental changes we expected. We do not use impact evaluations to consider all detailed aspects of the interventions, or to re-run a CBA.

2. We will generally use a mix of quantitative and qualitative tools. It is important that we do not only focus on those impacts that are most straightforward to quantify. Quantitative or qualitative evidence may reveal previously unknown, unintended consequences. We examine significant unintended consequences that might occur using the evidence gathered during an impact evaluation.

3. We aim to value the benefits of the intervention and relate these to the cost of the intervention, including through comparison with the expected costs and benefits in the CBA.

4. We usually only consider studies whose counterfactual is above a certain robustness threshold in systematic reviews that summarise the evidence on 'what works' in a particular sector. In some instances, studies just below that threshold (such as those included in Table A under constrained designs) can provide valuable lessons on our impact, if the limitations and caveats on any findings are made clear.

5. For our analysis, we generally start by using any regulatory data we already routinely collect from firms and existing consumer surveys, such as the Financial Lives Survey. Where this data is not sufficient for us to estimate the impact of an intervention, we consider what further data needs to be gathered and how frequently.

6. We also aim to understand the mechanisms by which the intervention has influenced the outcomes observed. We often include a process evaluation element, for example a check that remedies have been implemented as intended based on qualitative discussions with firms or consumers or our supervisors, when not obvious from the data.

7. Table A sets out a range of methodological approaches that help us measure with some confidence the extent to which we have caused any effects in the treated market.

Table A: Research designs and methods for use in evaluation

	Designs and methods	Description	Example of use
Unconstrained designs	Quasi experimental (QE) designs	Use a comparison group that is 'as good as random' or obtained in a way that allows us to model for selection bias. Various options for obtaining a comparison group to ensure the treated and untreated groups are as similar as possible.	When a policy can be introduced in a staggered fashion.
	Regression Discontinuity Design	Examines the boundary between the 'only just eligible' and the 'not quite eligible'. Results only apply directly to those at the boundary.	Used in EP23/1 to analyse the impact of repeat use remedies on customer outcomes.
	Difference in difference method	A method for analysing QE data. Compares how trends in outcomes change between treated and untreated groups over a time period relevant to the intervention. Unobserved factors might affect the outcome, but if they do not affect trends in the outcome then the trends for both groups in the absence of a policy will be the same.	Used in EP18/2 to identify the impact of our regulatory changes on the swap market. Tenures unaffected by the interventions were used as a control group.
	Interrupted time series design	No comparison group is available. The counterfactual is estimated from a projection of the outcome measure before the intervention.	When alternative causes for changes in outcomes can be eliminated and the impact is large compared with the error inherent in forecasting.
Constrained designs	Natural experiments and instrumental variables	Comparisons with a naturally occurring comparison group can be made even though none was present by design. Or an external factor can be identified, which influences the likelihood of being exposed to a policy and does not in itself affect outcomes.	When the policy has already been implemented and the opportunity to put an evaluation design in place at implementation was missed.
	Before and after studies	An outcome is measured before and after an intervention but there is no comparison group.	When the intervention is the only thing that could reasonably be expected to influence the result.

	Designs and methods	Description	Example of use
	Use of process evaluation information	Draws upon the findings of studies of the implementation and delivery of an intervention, often using qualitative methods, including case studies, e.g. front line staff often have a good feel for whether an intervention is effective or not.	When quantitative measures of impact are weak or not available. May capture a direction of change.

4 Case studies

1. The following case studies illustrate some of the key aspects of our approach to impact evaluation. All our impact evaluations are available on our website.

Combining quantitative and qualitative evidence

2. Combining quantitative and qualitative evidence help us identify not only what works but also why it works. It can also help with attributing any effects to an intervention when we cannot identify a comparison (non-treated) group. Box A provides an example of how we use both quantitative and qualitative analysis to evaluate our interventions.

Box A: Reducing barriers to entry into the UK banking sector

In 2018, we carried out an evaluation of a series of changes that had been made by us and the Prudential Regulation Authority (PRA) to the prudential and conduct requirements for new entrants to the banking sector. These changes had been recommended by the FSA and the Bank of England following a 2013 review. The intervention was aimed at ensuring that these requirements were proportionate and did not pose excessive barriers to entry.

We expected the intervention to make the authorisations process cheaper and quicker for potential new banks, reducing barriers to entry and increasing the number of firms entering the UK banking sector. We anticipated this would lead to an increased competitive challenge to existing banks and benefits to consumers across a range of products.

As there was no clear single and robust counterfactual for this intervention, we evaluated its impact by undertaking several discrete pieces of analysis to test key changes along the causal chain. This was done using a mix of qualitative and quantitative methods. Our econometric analyses included the following:

- using a difference-in-difference model to determine whether there had been an increase in licences issued because of our intervention, using 26 other EU countries as a comparator group
- using a regression technique called pooled ordinary least squares to compare the product offerings, including mortgages and fixed-term savings accounts, of post-review entrants to pre-review entrants and incumbents
- using descriptive statistics to compare the growth of deposit taking and lending activities of entrants before and after the review to see whether there had been any change

We conducted interviews and gathered qualitative insights from a sample of firms which had entered the retail banking sector since April 2013 to validate the outcomes we observed from the analytical approaches. The interviews focused on firms' experience of the authorisation process and the regulatory requirements since these changes were introduced, the extent to which it affected their decision to enter the market and how they sought to compete for custom within the retail banking segment.

By demonstrating that our expectations at the various stages of the causal chain had been met, we had some confidence, even in the absence of a single strong counterfactual, that any final positive impact was due at least in part to the 2013 intervention. This is typically known as theory-based evaluation.

Learning lessons to inform our regulatory approach

3. We share any lessons learnt through our evaluations so that our decision-making is informed by the evidence produced by impact evaluations. For example, lessons on the effectiveness of specific types of remedies informs future remedy design, alongside other market knowledge (see Box B).

Box B: Lessons learnt about outcomes-based remedy design from our overdrafts evaluation

The FCA's 2019 intervention into the overdrafts market included a requirement for firms to develop a strategy to reduce repeat use of overdrafts. Firms were given significant autonomy to decide how they would identify a customer as a repeat user, how and when they would reach out to them and what help to offer.

The policy was successful as it allowed firms to implement changes quickly. They were able to minimise their own costs and many felt ownership over improvements. Our impact evaluation found significant benefits to consumers helped by the strategies, saving them up to £177 in the 12 months after they entered the strategy.

However, there was variation in the size of the effects between firms, and in the number of consumers helped by the firms. The impact evaluation provided insight into the channels of communication that were effective, and examples of effective rules to identify repeat users. We have since published a 'good and poor practice' document.

Strengthening the links between CBA and evaluation

4. We have a legal obligation to undertake a CBA when consulting on rules unless an exemption applies. CBAs, as well as the analysis undertaken in market studies, help us provide a reliable baseline for impact evaluations. Strong links between CBA and impact evaluation make our impact evaluations as robust and effective as possible and help to ensure we have a continuous and viable programme of evaluations.

5. For example, an evaluation can identify significant differences between the value of costs and benefits and those estimated in the CBA. This evidence can help us make better assumptions in future CBAs. See also **Actions we can take after a review** in the Rule Review Framework for a discussion on the actions we may take after an impact evaluation.

6. Planning impact evaluation at early stages (eg prioritisation and CBA development) helps us in 3 key ways (see also):

    - First, it clarifies our thinking about expected outcomes that should be included in the CBA. Constructing a causal pathway is an effective way to visualise the

theory of change that underpins the CBA. This enables us to critically evaluate each link in the chain in a systematic way.

- Second, it can help ensure consistency in the CBA and impact evaluation. For example, it could help to align the 'without intervention' case in the CBA (often referred to as the 'do minimum') and the 'without intervention' (or counterfactual) case in the evaluation.

- Third, it can prompt us to store and collect the data needed for the evaluation. It is important we record the key assumptions underlying the CBA modelling. This will allow future evaluators to explore the causes of any discrepancies between outcomes in the impact evaluation and those expected in the CBA.

7. By identifying the key indicators needed for the evaluation against what data is available, evaluators can identify data gaps that require bespoke data collection, including baseline data, and monitoring. This must be done at the CBA stage because any additional data must be collected before implementation, and in some cases before we publish the consultation paper. This is due to potential 'anticipation' effects, such as prices or sales strategies reacting immediately to the announcement. Our consultation on Rent-to-Own (CP18/35) signalled our intention to continue to collect data to enable us to evaluate such intervention in the future, which we did in December 2020 (Evaluation of our Rent-to- Own price cap).

8. Box C provides an example of an intervention where the impact evaluation mirrors closely the CBA, while attempting to provide additional information on the relative outcomes of 2 remedies.

Box C: Guaranteed Asset Protection

In July 2014, we published the general insurance add-on products market study final report. The market study identified significant consumer harm in the sale of add-on Guaranteed Asset Protection (GAP) insurance products, including around £76-121 million in overpayments (out of an estimated market size of £152 million).

This overpayment was attributed to the point-of-sale advantage held by add-on sellers, as well as a lack of consumer understanding of GAP insurance products. To address this, we intervened in 2015 by making it mandatory for vehicle sellers to provide sufficient information to consumers and requiring a 2-day pause in the sale before it could be concluded ('deferred opt-in'). We believed that having both time and information would enable consumers to decide whether they need GAP insurance, and to shop around if they do.

We expected:

- improved competition between add-on and standalone sellers
- better outcomes during the purchasing process, including an overall decrease in add-on GAP insurance sales, given our concern about consumers buying unsuitable add-on products
- more consumers shopping around and buying GAP insurance from standalone providers

To assess the extent to which these impacts occurred, and if these impacts arose due to our intervention, our impact evaluation looked at the change in the following indicators, many of which were considered in the CBA:

- the price of add-on GAP insurance
- the quantity of add-on GAP insurance sold
- the marginal cost of supplying add-on GAP insurance
- the share of add-on GAP insurance sales to total GAP sales and to car sales
- the share of GAP insurance sales between add-on and standalone providers
- consumers' engagement with the process of buying GAP insurance (eg shopping around)

Our report included the following key elements:

- work to understand what had happened in the market since our intervention through insight gathered from discussions with industry trade bodies and an analysis of market-level industry data
- an econometric analysis of firms' data to diagnose and isolate the impact of our interventions on the GAP insurance market
- a commissioned survey of those consumers who had recently purchased cars, and so had the opportunity to buy GAP insurance following our intervention.

5 Key challenges

1. Impact evaluation can be complex and challenging. Sometimes findings from evaluations are mixed or unclear. There may be uncertainty around the quantified net benefits, there may be positive elements of interventions even when they have not worked as anticipated overall or the impact of some interventions simply cannot be evaluated with any degree of robustness. In this section, we explain some of the key challenges for impact evaluations and how we address them.

Time since the intervention

2. There is often a lag between our activities and resulting market changes. For example, it may take time for market study remedies to be implemented as rules, for compliance systems to be put in place and for identifiable changes in behaviour to take effect (although some behaviours might change in anticipation of rule changes). Even after implementation, some effects might happen immediately, while others will take longer. We may only expect the full benefits of an intervention to come about later.

3. Undertaking an evaluation too soon may lead to no impact being identified and an inability to quantify any benefits to consumers, even though these materialise later. On the other hand, waiting too long before evaluating an intervention:

    - makes it too difficult to attribute impact to the intervention, as other factors also affect the market
    - may lead to memory recall problems when asking firms and consumers about the intervention
    - delays remedial action to address shortcomings in the intervention

4. If we are interested in the exact impact of a measure, we may look at impacts right around the intervention date so that we do not conflate with broader trends and dynamic responses. However, if the remedy was introduced with dynamic interactions between firms and consumers in mind, we may be more interested in looking at the outcome after enough time has passed for these effects to have taken place. Both are valid approaches in evaluation.

5. Planning for evaluation at an early stage helps to address this timing challenge. For example, by collecting any necessary data close to implementation, and then undertaking the final impact evaluation assessment at the appropriate later stage.

Quantifying and attributing impact to the FCA

6. The counterfactual may be difficult to identify (see, for example, our evaluation of reducing barriers to entry in banking). This makes impact difficult to quantify, as markets may change over time due to factors external to the FCA, for example innovation. To estimate impact correctly it is crucial to establish the right counterfactual. For example, a lack of positive change in outcome measures after an intervention may not be a bad result. There might have been a negative change if

the intervention had not taken place. Other regulatory changes happening around the same time also need to be considered.

7. There may be instances where we do not have a strong counterfactual, or where this cannot be identified in advance. This is inevitable because we want to focus on evaluating areas that matter, not just those where the counterfactual is most robust. Intervening events (for example actions by other agencies) can change market conditions significantly. This can mean it is no longer possible to isolate the impact of a specific intervention and in some cases we may not be able to carry out an intended impact evaluation of interventions.

8. In other cases, we may judge it is important to go ahead with an evaluation because we can learn from it even if the counterfactual is not as robust as we would like. This may include studies where it can be difficult to identify a counterfactual, but data may nevertheless be revealing. For example where a market shrinks or the price jumps post-regulation as theory would predict. To come to a view about causality, those 'before and after' studies will typically consider intermediate as well as final outcomes and triangulate from a variety of sources to consider if other factors have influenced the results or data. Other regulators and agencies, such as the CMA, have used similar techniques in the past. See Box 1 in the Rule Review Framework for information on approaches designed to reduce uncertainty about the contribution the intervention is making to the observed results.

9. Sometimes we need to accept that it may not be possible to fully quantify the impact downstream of an intervention upstream. In addition, some outcome measures might be difficult to quantify (e.g. quality, innovation) and we need to rely on proxies.

10. Very often, our interventions include a package of inter-related remedies, rather than a single remedy, and we need to decide whether we want to understand individual or combined effects. For example, see Box D on the sale of add-on Guaranteed Asset Protection (GAP) insurance products. In the evaluation, we tried to learn about the relative efficacy of 2 remedies. In other instances, we may only be interested in the combined effect.

11. In some cases, there will be various actors involved (eg the Treasury, the Bank of England or the CMA). It may be hard to separate the work of the FCA from that of others (for example, our evaluation of the impact of bringing additional benchmarks into the regulatory and supervisory regime). In those instances, we can review particular outcomes of an intervention rather than attempt to measure every aspect or the whole intervention. For example, we may decide to focus on whether there are signs of improved conduct and whether the market is 'cleaner', rather than attempting to disentangle the effect of each policy requirement on firms and consumers, or to separate contributions of the FCA from other organisations, if it is reasonable to suppose that our intervention had a part to play.

Ensuring robustness and credibility

12. To be credible, evaluation work needs to be robust and done with sufficient independence. One way to ensure robustness is to have strong links between CBA and impact evaluation.

13. We are committed to independent impact evaluations. We may either externally commission them (or any individual elements) or, when we undertake them internally building on our knowledge and expertise, we seek input from objective external experts. They peer review the method and the quality of evaluations. The impact evaluations we have published so far have all benefited from the advice and peer review of academics who are experts in evaluation techniques and/or the sectors we were evaluating. In some cases, we have also sought input from more than one academic. We also ensure any impact evaluation will be led by a different team from the one that introduced the intervention being evaluated. We will continue to consider further ways to ensure objectivity in our impact evaluations.