

**Proposed Framework for Post-Implementation Evaluation of the
Effects of the G20 Financial Regulatory Reforms**

Consultation document on main elements

11 April 2017

Contents **Page**

- 1. Introduction 3
- 2. Discussion questions 4
- 3. Objectives and scope of the framework 5
- 4. Concepts and terms 6
- Figure 1: Stylised Policy Evaluation Concept..... 10
- 5. Evaluation approaches..... 10
- Table 1: Categories of Tools and Types of Evaluations 13
- 6. Responsibilities and policy evaluation process 14

Contacting the Financial Stability Board
Sign up for e-mail alerts: www.fsb.org/emailalert
Follow the FSB on Twitter: [@FinStbBoard](https://twitter.com/FinStbBoard)
E-mail the FSB at: fsb@fsb.org

Proposed Framework for Post-Implementation Evaluation of the Effects of the G20 Financial Regulatory Reforms

Consultation document on main elements

1. Introduction

The G20 launched a comprehensive programme of financial reforms post-crisis to increase the resilience of the global financial system, while preserving its open and integrated structure.¹ By making the financial system more resilient and thereby reducing the likelihood and severity of crises, the reforms support the G20 objective of strong, sustainable and balanced growth. The role of the Financial Stability Board (FSB) has been to coordinate the development and support the full, timely and consistent implementation of these reforms.

With the main elements of the post-crisis reforms agreed and implementation of many core reforms underway, initial analysis of the effects of those reforms is becoming possible. Implementation monitoring and the evaluation of the effects of reforms represent good regulatory practice, form part of the FSB's accountability to the G20 and the public, and inform structured policy discussions among FSB members and standard-setting bodies (SSBs).

As stated in its second annual report to the G20 on the [*Implementation and Effects of the G20 Financial Regulatory Reforms*](#), the FSB is working to enhance the analysis of the effects of those reforms, including whether they are working together as intended. To that end, the FSB, in close collaboration with the SSBs, and informed by work carried out by its members and other stakeholders, is developing a structured framework for the post-implementation evaluation of the effects of the G20 financial regulatory reforms ("the framework") during the German G20 Presidency.

Building on the experience with the FSB's 2011 [*Coordination Framework for Implementation Monitoring*](#) (CFIM), the policy evaluation framework will specify processes and analytical approaches for the evaluation of reforms. The G20 Finance Ministers and Central Bank Governors, at their March 2017 meeting in Baden-Baden, welcomed the FSB's work to develop the framework. Application of the framework will begin over the coming years, with incremental improvements to the framework being added as appropriate. The purpose of this consultation document is to seek the views of stakeholders on the main elements of this framework; its scope, concepts, methodologies and approaches. Moreover, the public consultation provides an opportunity for respondents to highlight useful empirical tools, as well as challenges inherent in evaluations. Following this public consultation, the framework will be presented to the Leaders' Summit in Hamburg in July 2017 and published.

¹ For more information on the G20 financial regulatory reform agenda, see, for example, the [FSB's progress reports to the G20](#) and the second annual report on the [*Implementation and Effects of the G20 Financial Regulatory Reforms*](#).

2. Discussion questions

The FSB invites feedback on the following questions:

Main elements of the framework:

1. Do you have any comments or suggestions on the main elements of the evaluation framework (e.g. are there other elements that should be considered for inclusion in the framework)?
2. Are the objectives and scope of the framework appropriately set out?
3. Would you suggest any refinements or additions to the concepts and terms?

Challenges of evaluations:

4. Do you have comments or suggestions on how to address the challenges of identifying and measuring interactions between reforms and how to isolate the effects of reforms and their interactions from other factors?
5. Do you have views on how to think about intended versus unintended (and possibly undesirable) consequences or how to frame the trade-off between different (and possibly competing) objectives?
6. Do you have comments or suggestions on how to address the challenges of defining and measuring social benefits and costs, especially when they do not follow directly from private benefits and costs?

Evaluation approaches:

7. Do you have comments or suggestions on the proposed evaluation approaches (i.e. on the empirical models and methods to analyse effects)?
8. Do you have suggestions on approaches to ensure the quality and replicability of results?
9. Do you have views on lessons – in terms of methods and approaches – that can be learned from evaluations in other policy areas, or from existing national or regional evaluation frameworks?

Data issues:

10. Do you have suggestions on information sharing arrangements (publication of results, repository of evaluations, and data availability, particularly as it pertains to replicability)?

Engagement with stakeholders:

11. How can the FSB and SSBs best engage with external stakeholders (e.g. financial services providers, various kinds of end-users, and academics) in their evaluation work (going beyond public consultations)?

Prioritisation of topics:

12. Do you have comments or suggestions on which individual reforms or interacting set(s) of reforms should be initially considered for evaluation as a matter of priority?

3. Objectives and scope of the framework

Objectives of the framework

The framework will serve as a guide for post-implementation evaluation work by the FSB and SSBs. In particular, it aims to guide analyses of whether the G20 financial regulatory reforms are achieving their intended outcomes, and help identify any material unintended consequences that may have to be addressed, without compromising on the objectives of the reforms. Applying such a framework will inform structured policy discussions among FSB members and SSBs. Evaluations, if findings warrant it, could provide a basis for possible fine-tuning of post-crisis regulatory reforms, without implying a scaling back of those reforms or undermining members' commitment to implement them.

The regulatory reform process follows the policy cycle, which comprises: (i) identifying a market failure, risk or problem that requires attention; (ii) designing, consulting on, and adopting the appropriate policy response; (iii) implementing the policy and monitoring that implementation; and (iv) evaluating the effects, and where warranted, revisiting step (ii) to determine whether any additional steps are required in light of sufficient evidence. The length of the policy cycle would differ across reform areas, not least because different regulatory measures can take more or less time to play out through the financial system and the real economy. A structured evaluation framework adds to the policy cycle by setting out general guidelines for post-implementation evaluations, by clarifying roles and responsibilities, and by helping to select and prioritise individual evaluation projects.

As such, the framework should help to manage the methodological challenges inherent in any policy evaluation study. These challenges include analysing a multitude of transmission channels and behavioural responses, accounting for complex interactions among regulatory and non-regulatory factors, estimating benefits and costs (including long-term effects and social costs) and considering data gaps (see Box 1).

These challenges are not unique to financial reforms, and are relevant for policy development and evaluation in other areas (e.g. competition, tax, environmental, health, and labour-market policies). Addressing these challenges through robust evaluations will evolve with experience. Thus, the framework is a living document that will be enhanced as experience is gained and its roll-out will take place progressively in the coming years.

Scope of the framework

The focus of the framework is on post-implementation evaluation; that is, evaluating the effects of G20 reforms for which implementation is well underway or completed. The framework will cover ex ante analysis only to the extent that it is necessary for post-implementation evaluation, such as translating policy objectives, set out at the time of development, into observable and measurable outcomes, and comparing available ex ante impact assessments with realised outcomes.

The framework is being designed with the post-crisis G20 core reform areas in mind, but it should provide flexibility to allow tailoring evaluations to other – including future – reforms, and the potential for individual jurisdictions to use it, if they desire. In particular, the core reform areas to which the framework will be applied are: (1) making financial institutions more resilient; (2) ending too-big-to-fail; (3) making derivatives markets safer; and (4) transforming shadow banking into resilient market-based finance. The framework applies both to individual reforms,

as well as to cross-sectoral and cross-cutting issues stemming from the interaction and combined effects of those reforms, on key aspects relating to the functioning of the global financial system.

4. Concepts and terms

An important objective of the policy evaluation framework is to clarify relevant concepts and terms, with the aim of achieving a common understanding by FSB members and relevant stakeholders.

Implementation monitoring

Over the past few years, the FSB and SSBs have established various implementation monitoring and reporting mechanisms in accordance with the CFIM. These mechanisms² have focused on the timeliness of implementation by member authorities (via laws, regulations, and supervisory guidance) and, where possible, on the completeness and consistency of domestic rules with international standards. Knowledge gained from monitoring efforts to examine whether and to what extent the reforms have been implemented in a comprehensive, timely and consistent manner will be key inputs into the reform evaluation process. Equally, policy evaluations may be useful inputs to the identification of gaps in implementation monitoring and/or in addressing key risks.

Evaluation objectives

The starting point for any post-implementation evaluation exercise is to set out the reform's original objectives and the primary issues that the reform was intended to address. This latter step involves, among others, defining the underlying market failure(s) or negative externality(ies) that were identified at the time of the reform development. *Broad objectives* are set out by the G20 relating to a resilient, open and integrated global financial system that supports strong, sustainable and balanced economic growth. *Specific objectives* are set out by the relevant body(ies) for individual reforms.

Policy evaluations require the articulation of measurable outcomes (operationalisation) that can be compared against benchmarks distilled from these objectives. Since a reform's objectives are often expressed in fairly general (non-technical) terms, they must be translated into *measurable benchmarks* to assess whether the reform's effects are in line with what was desired given the underlying ex ante assumptions.

Evaluations in a given reform area are not expected to be repeated frequently. Monitoring developments in the relevant market segment, e.g. through a set of indicators updated on a regular basis, would help to identify issues and trends (including any regulatory gaps, remaining or newly identified benefits or risks, or potential positive or negative unintended consequences) that could subsequently be examined in detail during evaluations. Such metrics should be developed and reviewed periodically by the bodies responsible for evaluations.

² Such as the Basel Committee's Regulatory Consistency Assessment Programme, similar exercises conducted by the Committee on Payments and Market Infrastructures (CPMI) and the International Organization for Securities Commissions (IOSCO), and the FSB's Peer Reviews.

The monitoring of selected metrics is a precursor to evaluations and is useful for reviewing trends and identifying issues on a regular basis. Once deemed necessary, full evaluations require more complex analyses.

Effectiveness of individual reforms, interactions and overall effects of reforms

The evaluations undertaken by the FSB and SSBs may include analyses from three distinct but inter-related perspectives:

- Evaluation of the effectiveness of *individual reforms*
 - The effectiveness of individual reforms may be evaluated by comparing outcomes with the specific reform objectives, typically outlined in the policy document by the relevant SSB or the FSB. One such example would be to compare observable outcomes with the Basel Committee on Banking Supervision's (BCBS) liquidity coverage ratio's objective "to promote short-term resilience of a bank's liquidity risk profile by ensuring that it has sufficient high quality liquid resources to survive an acute stress scenario lasting for one month."³
- Evaluation of the *interaction and coherence* among reforms
 - This type of evaluation would examine whether the combination of particular reforms has reinforced or impeded the attainment of their respective specific reform objectives or led to other (e.g. beneficial or adverse) effects. An example of possible interactions among reforms is the Basel III leverage ratio and central clearing requirements for over-the-counter (OTC) derivatives. It is important to analyse how reforms interact (e.g. where or how they may reinforce each other or whether they might work at cross-purposes), and whether individual reforms are appropriately calibrated once their combined effects are taken into account.
- Evaluation of *overall* effects
 - This type of evaluation (the most challenging of the three) would examine the extent to which a particular reform, individually as well as in combination with other reforms, has contributed to the broad G20 objectives of strong, sustainable and balanced growth. From the perspective of the global financial system, these objectives relate primarily to financial system resilience, the orderly functioning of markets, and the cost and availability of financing to households and nonfinancial businesses. The need for evaluating overall effects may also arise from the identification of cross-cutting issues (e.g. concerns about impaired market liquidity or the shift of activities outside the regulatory perimeter) that may involve the confluence of different reforms and other factors.

Measuring benefits and costs

Evaluating the overall effects of reforms involves the estimation of *benefits and costs*. Such evaluations should focus on assessing the *social* benefits and costs. They should consider

³ Basel Committee on Banking Supervision, *Basel III: A global regulatory framework for more resilient banks and banking systems*, December 2010 (rev June 2011), www.bis.org/publ/bcbs189.htm.

private benefits and costs that accrue to particular market participants or end users to help assess social benefits and costs and shed light on relevant distributional issues. It is also important to distinguish between temporary and permanent effects, since some of the costs of the regulatory reforms may only be applicable during the transition period.

Commonly identified social benefits include lower probability and severity of financial crises and their associated output losses, reductions in funding advantages and related distortions owing to perceptions that some institutions are too-big-to-fail, and improved resource allocation owing to reduced financial and economic cyclicality. Other benefits may also be considered, e.g. in terms of reducing market abuse, enhancing transparency for end users as well as enhancing innovation and efficiency.

Commonly identified private costs include adjustment costs by market participants to comply with new rules (e.g. in terms of increased staff and enhanced information technology systems), while social costs often relate to potential knock-on effects on economic growth caused by possible (temporary or permanent) reductions in the availability of financial intermediation.

Benefits and costs may arise from the achievement of the *intended* objectives of the reforms; and from unintended consequences, which can be positive or negative. Additionally, benefits or costs could arise if, for example, the reforms lead to fewer or create new sources of risk for the financial system. In considering social benefits and costs, evaluations should, ideally, also consider any *trade-offs* in attaining different reform objectives, such as improvements in financial system resilience that negatively impact financial intermediation (or vice versa). Where possible, evaluations should also consider *distributional* effects (e.g. allocation of benefits and costs across jurisdictions, sectors and end users).

While some of the costs of reforms for market participants may be measurable in the short term, measuring the accreting overall social benefits (e.g. in terms of crises avoided or tempered) is far more difficult. More broadly, the effects of implemented reforms can only be fully ascertained over a longer period of time that includes a full financial cycle and both stressed and normal market conditions.

The analysis of benefits and costs may also involve, at least to a certain degree, a consideration of *efficiency*. Efficiency relates to the question whether reforms are working in a way that minimises net social cost for a pursued social benefit, helping both policy makers and stakeholders to draw conclusions on whether the costs of an intervention are proportionate to the benefits.

Stylised policy evaluation concept

Assessing the effects of the G20 reforms will require a sequence of evaluations complementing each other. This policy evaluation concept involves a number of important considerations, including:

- Timing and frequency of evaluations
- Priority of evaluations
- Scope and depth of analysis
- Available data and information

Figure 1 below illustrates, in a stylised manner, how these critical considerations fit together.

Along the x-axis is the timeline of the policy/reform implementation. The blue area, to the left of the y-axis, represents the policy/reform development stage. Once the reform is put in place, implementation begins, but its timing will vary by jurisdiction (thus the shaded blue-green area, representing a transition). The green area (the post-implementation period for nearly all jurisdictions) represents the period when evaluations could be conducted (reflecting that the focus of this analysis applies to *post-implementation* evaluation). Evaluations only become feasible once a critical mass of jurisdictions have implemented the reform and a sufficient amount of data has become available. The timing of evaluations should also consider the point in a financial cycle, given that the effects of implemented reforms can only be fully ascertained over a longer period of time that includes a full financial cycle, with both stressed and normal market conditions.

Prioritisation of evaluations

Along the y-axis is the priority given to a specific evaluation. Prioritisation is one of the most important considerations. Having a transparent system for the *prioritisation* of evaluations is instrumental to appropriately plan and allocate limited resources, and in view of different implementation schedules (among the reforms and jurisdictions). While the framework is intended to be applicable to all post-crisis G20 reforms, it should help determine whether, when, how frequently and at which analytical depth to conduct evaluations of individual reforms or groups of reforms.

Decisions on the priority of evaluations can be made on the basis of *materiality* and *feasibility*. Materiality is the relative importance of the reform or cross-cutting issue. Feasibility is the capacity to carry out a meaningful evaluation, which is conditional on a variety of factors (see below). Materiality should take into consideration *the potential effect of the observed reform or cross-cutting issue* on global financial stability and other objectives set by the G20 (e.g. orderly functioning of markets, costs and availability of financing, financial integration), and also the *defined priority of the reform* (e.g. based on the distinction between core reforms and other reforms that FSB members have already agreed upon for policy and implementation monitoring purposes).

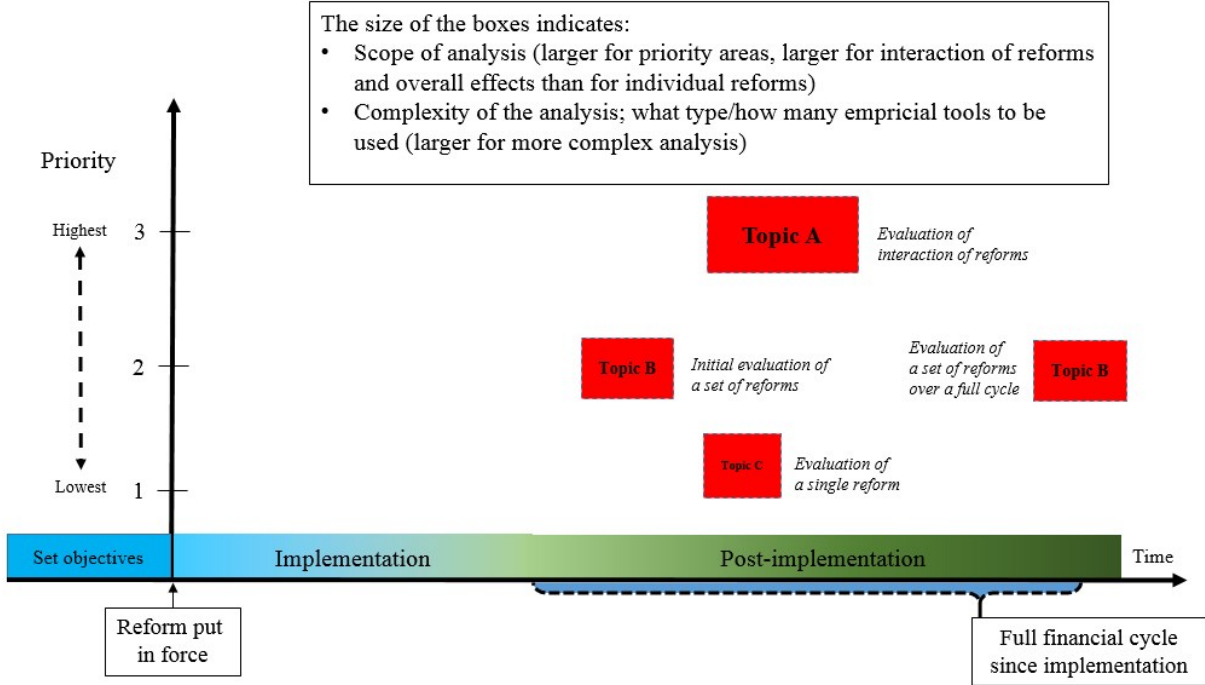
The actual timing, duration, and frequency of evaluations would also be determined by *feasibility* factors such as *implementation progress* (the extent to which a given reform has been implemented and is consistent with the international standard, where relevant, and with respect to regulatory outcomes),⁴ *choice of objectives* to evaluate (specific or broad), *availability of resources* to carry out evaluation work, desired complexity of the evaluation, and allowance of sufficient time for *relevant data* to become available.

The size of the boxes in the Figure indicates the scope and complexity of the analyses; larger boxes representing more complex and broader scope analyses. Evaluations will differ by

⁴ This information is collected by the SSBs and FSB in some reform areas, and is summarised in the annual FSB report to the G20. For example, the BCBS monitors the adoption of Basel III standards on a semi-annual basis (formerly “level 1 assessments”); carries out jurisdiction assessments on the extent to which domestic regulations are aligned with the minimum Basel requirements (formerly “level 2 assessments”); and examines bank implementation of the Basel requirements, and assesses whether prudential ratios are calculated consistently by banks across jurisdictions to improve comparability across outcomes (formerly “level 3 assessments”). CPMI and IOSCO follow a similar approach for monitoring implementation of some of their respective standards.

reform, with the choice of tools affected by the complexity of the topic. The evaluations will also differ by type (individual reforms, interactions or overall effects of reforms), and available data and information. Topic A (the top box in the Figure) represents the evaluation of a cross-cutting issue, and is high on the priority scale. Topic A is larger than the others, reflecting the complexity of that particular issue. Topic B is the evaluation of a set of reforms, and the second box B signals that the analysis will be repeated at a future point in time (after a full financial cycle since implementation). Topic C, the smallest box in the Figure, signifies that this evaluation, of a single reform, is lowest on priority between the three and the least complex. More generally, given these and other factors, the framework must allow for a flexible design of the evaluations and cannot have a “one-size-fits-all” approach.

Figure 1: Stylised Policy Evaluation Concept



5. Evaluation approaches

In designing policy evaluations, it is important to consider their inherent conceptual challenges. In particular, evaluations of post-crisis reforms should address three main questions:

- (1) Did the reform “cause” an outcome? (*Attribution*)
- (2) Did the reform have broadly similar effects across relevant markets, states of world, or jurisdictions and regions? (*Heterogeneity*)
- (3) Did the reform achieve its overall objective? (*Aggregation/general equilibrium*)

Aiming to address these three questions (further detailed in Box 1), which are not exhaustive and have significant overlap, is fundamental for any proper policy evaluation. To answer the question of whether a reform “caused” an outcome (*attribution*) requires isolating its effects from other policies and factors, including conjunctural ones, and selecting appropriate counterfactuals and baselines. It also means considering both the reforms’ short-term

adjustment costs for the financial industry and end users, generally easier to measure, and their long-term benefits obtained, often more difficult to quantify in deterministic terms (e.g. crises avoided or tempered). Analyses need to consider that reform effects can differ given varying institutional-, market-, legal-, and jurisdiction-specific structural settings and possible differences in implementation across jurisdictions (*heterogeneity*). To assess whether a reform achieved its overall objective (*aggregation/general equilibrium*) requires one to consider the many channels by which it is transmitted to the financial system and real economy, the various behavioural responses and feedback effects (e.g. macro-financial linkages), and complex interactions. Crucially, as mentioned above, overall evaluations need to assess not just the private *benefits and costs* to market participants and end users, but the social *benefits and costs*, which is complex, given the possible trade-offs involved in reform objectives, the still incomplete financial intermediation theory and data gaps.

Box 1: Conceptual challenges to policy evaluations

Attribution: Did the reform “cause” an outcome?

- Selecting appropriate *benchmarks* and *counterfactuals* for evaluating outcomes is important, but challenging. The appropriate baseline to use is not always clear, e.g. many pre-crisis indicators are likely not appropriate benchmarks from a stability standpoint, since many of the financial circumstances and behaviours proved unsustainable.
- Financial stability outcomes are driven by the behaviour of financial institutions and other financial markets participants, which, in turn, are affected by other policy actions and exogenous factors, in addition to the reforms. Thus, it is often difficult to identify an appropriate counterfactual scenario that would have materialised in the absence of a regulatory standard and to single out the post-implementation effects of individual reforms from other policies and factors, including conjunctural ones (e.g. unconventional monetary policies).
- Many core reforms are still being implemented, so longer-term (steady state) effects including, in particular, social benefits may not yet be apparent or measurable and observed trends may be temporary in nature. Related, short-term adjustment costs for the financial industry and end users are generally easier to measure than long-term accreting benefits (e.g. crises avoided or tempered).
- Reforms are often phased in considering the state of the financial sector. Thus, the rollout of the reforms may be shaped by effects experienced to date, confounding causal interpretations.

Heterogeneity: Did the reform have similar effects across markets, states of the world, or jurisdictions and regions?

- How reforms are implemented in individual jurisdictions may vary – contributing to differences in effects – yet indicators of the quality and intensity of reforms can be hard to develop, reducing the ability (statistical power) of the analyses to detect these differences. Moreover, how reforms are implemented in other markets and jurisdictions may affect outcomes, e.g. as they create spillovers, that can confound the analyses.
- Effects can differ across markets, states of the world or jurisdictions and regions given different starting positions, institutional and structural settings, and scopes for substitution among financial providers and activities.

Aggregation/general equilibrium: Did the reform achieve its overall objective?

- Financial intermediation theory is incomplete, particularly with respect to capturing the quantitative, general equilibrium effects of financial intermediation on the real economy, and

vice-versa. As the set of drivers of financial vulnerabilities and systemic risk, as well as the transmission channels and aggregation of effects, are not yet well identified in existing models, they may not be fully captured in empirical analyses.

- Linking micro and sectoral analyses and indicators of progress with macro-level objectives is both conceptually and empirically difficult, given the many factors shaping aggregate conditions, including feedback effects (e.g. real-financial linkages, effects of reforms on asset prices) and complex market interactions and related externalities. This aggregation is compounded by the lack, at this stage, of adequate post-implementation data, and more generally, of data gaps.
- The ultimate objective of the evaluations will be to assess social benefits and costs. By definition, these are particularly difficult to capture and measure.

Evaluation tools

It is important to match the available evaluation tools with their ability to address the above questions (on *attribution*, *heterogeneity*, and *aggregation/general equilibrium*). While they arise in all three types of evaluations – effectiveness of *individual reforms*, *interaction and coherence* among multiple reforms, and *overall* effects – addressing these questions may be harder for some types of evaluations than for others. And the preferred tool may depend on the specific reform, the timing and priority of the evaluations, the desired scope and depth of the evaluation, and the availability of data and analytical methods. As such, the preferred tools should vary by both the specific type of evaluation and the reform being evaluated.

The tools identified as being useful for policy evaluations, and increasingly used by policy makers and researchers, range from qualitative analyses to simple and more complex empirical methods, grouped here in four categories:

1. *Qualitative analyses* include peer reviews and other exercises that rely on questionnaires and assessment methodologies or benchmarks to evaluate a reform.
2. *Indicators and descriptive statistics* include indicators and descriptive statistics of the sort often used in financial stability reports.
3. *Partial equilibrium analyses* can be grouped into event studies, regression techniques and, broadly, other techniques, with a theoretical or empirical focus but not capturing overall effects.
4. *General equilibrium analyses* include time-series models with specific structures and comprise theoretical and calibrated methods as well as empirical techniques.

To provide guidance on the specific tools that may be useful for a given analysis requires assessment of how well each tool addresses the three questions listed above. Table 1 presents the four categories of tools (the left most column) and the three types of evaluations (the top row), along with the essential step of implementation monitoring/intermediate metrics (second to the left most column).

Table 1: Categories of Tools and Types of Evaluations

Category of tools	Type of evaluation	Implementation monitoring (pre-evaluation analysis)	Evaluation of individual reforms	Evaluation of interactions and coherence of reforms	Evaluation of overall effects
	1. Qualitative analysis	A, H	A, H	H	H
	2. Indicators and descriptive statistics	A, H	A, H	H	H
	3. Partial equilibrium type analysis		A*, H*	A, H	
	4. General equilibrium type analysis			A, H, G	A, H, G*

Notes: Attribution = A, Heterogeneity = H, General Equilibrium = G; a * after one of the alphabetic codes signifies that the tool is particularly useful. Grey shading signifies that the tool is not particularly useful for this type of evaluation.

Choosing a set of tools

The scoring of tools, in terms of their strengths and weaknesses, is depicted in each cell of Table 1. It shows that tools can vary in their ability to address the three questions outlined above (e.g. some tools can address attribution and heterogeneity better than others, but may be less useful to aggregate results).

As methods vary greatly in their ability to address the challenges, there is no “one-size-fits-all” approach. There is, however, a progression in the use of methods. Qualitative analysis and indicators will often be the starting point; using simpler indicators and descriptive statistics may be particularly useful to provide some insights for subsequent, more comprehensive and complex evaluations, such as those encompassing multiple reforms implemented independently or at different times. Evaluations to identify causal relationships and assess overall impacts will typically call for more complex analyses, with greater requirements for data and specialised expertise.

In practice, there may be trade-offs between some of the tools and the number and choice of tools may depend on the granularity and intensity of the assessments. And the choice of the appropriate tool(s) may be limited by data availability. However, multiple tools will likely be called upon for most evaluations as they complement each other. Using several methods can enhance the robustness and reliability of evaluations. Most evaluations of major reforms and of the interaction of reforms could be expected to employ, at some point, the most complex types of analyses, including general equilibrium analyses. For some reforms, however, it may not be necessary to consider general equilibrium effects and, for some reforms, such analyses may be infeasible given the available data and techniques. General equilibrium models are difficult to develop, but, in principle, they allow for considerations of the relevant transmission channels and feedback mechanisms that other methods do not allow. All methods have their limitations, and evaluations should explicitly state the limitations, simplification and assumptions taken to appropriately set expectations on what evaluations are able to accomplish.

6. Responsibilities and policy evaluation process

The FSB's mandate to coordinate at the international level the development and implementation of effective regulatory, supervisory and other financial sector policies places a responsibility on the FSB, together with the SSBs, to assess the overall coherence and consistency of evaluations across reform areas and to coordinate plans where needed.

Responsibility for policy evaluations

The relevant SSBs and the FSB will take primary responsibility for evaluations in reform areas where they have issued the relevant policies or regulatory standards. The FSB and the relevant SSBs will coordinate on interactions, coherence and overall effects of reforms, drawing on the coordination structure agreed upon in the CFIM. Evaluations would build as much as possible on SSBs' existing implementation monitoring and assessment frameworks, and would be carried out in accordance with those bodies' governance structures. In general, these more complex types of analyses would build on evaluations of individual reforms carried out by the SSBs and the FSB.

Resourcing of evaluations

Policy evaluations are resource-intensive exercises. Having a structured framework will help to promote effective allocation of resources. Prior to any decision to start an evaluation, resources should be clearly identified and evaluation exercises be prioritised accordingly.

An SSB or the FSB may elect to delegate the exercise, e.g. to a task force comprising representatives of member jurisdictions, while maintaining primary responsibility for the exercise. Because evaluations should seek to be as rigorous as possible, evaluation teams should comprise staff with relevant policy experience, as well as technical skills in economic analysis, research design and quantitative modelling. Consideration should also be given to the management structure of the evaluation team.

Design of the policy evaluation process

The *design of the evaluation process* will depend on the specific reform area or cross-cutting issue to be evaluated and the allocation of responsibilities among different bodies, taking into account the criteria and considerations set out in this document. Evaluations should be designed to be as robust as possible, but should also acknowledge the practical limitations, in particular, with regard to data access, quality and availability, inherent difficulties in selecting appropriate benchmarks and counterfactuals for evaluating outcomes, and methodological challenges, as outlined in the Box. The evaluations will include processes that ensure quality and objectivity of results.

The evaluation process should facilitate appropriate consultation and collaboration between all bodies involved, including other stakeholders where appropriate, and the public with respect to the outcome of the evaluation. On that basis, the design of the evaluation process should consider the following elements:

- Data and information requirements
- Engagement with stakeholders
- Transparency
- Follow-up

The *data and information requirements* for evaluations would be determined on a case-by-case basis by the body in charge of the evaluation, taking on board comments from the FSB and SSBs. Issues related to data confidentiality and other data usage restrictions, data quality and replicability, as well as costs and time involved if new data collection is required or desired, should be addressed sufficiently early in the evaluation planning process. Data availability, access and quality are likely to limit the types of evaluations feasible. The ongoing work under the G20 initiatives of closing data gaps and improving data sharing may provide some helpful lessons going forward.

The evaluation process should also include means for input from and *engagement with stakeholders* in the design of evaluations, the collection of relevant information, and the conduct of the analysis, thereby leveraging expertise from outside experts, including academics, think-tanks, industry and others. This engagement may include events such as workshops or calls for evidence. In accordance with the governance protocols and confidentiality arrangements of the relevant bodies, consultation processes and content may vary across SSBs and the FSB.

On collection of relevant experience and analysis, FSB members have suggested that a *repository* of relevant evaluation studies could be helpful. Similar repositories exist in other areas such as medicine. The repository would be accessible by a broad base of stakeholders and support the standardised and centralised collection of information on the effects of reforms. Its main function would be to collect and make publicly available evaluation studies and other analyses that are relevant to the G20 reforms (e.g. academic papers, studies by official sector, or by the financial industry). Whenever possible, data used in the analyses should be also made public.

Transparency: The FSB, in collaboration with the relevant SSBs, will discuss the results of the evaluations for each area and determine issues to highlight and relevant material to include in its annual report to the G20. For transparency and accountability, there will be an expectation to seek public comments on the outcome of the evaluations and to make the information in the evaluation reports publicly available, subject to confidentiality frameworks, specific arrangements and FSB members' consent. Decisions on publication will be made on a case-by-case basis, based on the recommendations of the body(ies) responsible for the evaluation.

Follow-up: The results of the evaluations will not contain any specific policy recommendations, but rather findings for consideration by the appropriate bodies. In accordance with established processes in the SSBs and the FSB, results of evaluation reports will be sent to the relevant body for further consideration of their implications for a given reform. The final responsibility for deciding whether and what amendment to a particular reform to propose to the G20 reform remains with the relevant SSB or the FSB.

Process for refining the framework

The framework will be refined as experience is gained. The FSB, in close collaboration with SSBs and other stakeholders, will coordinate this work.