

# Evaluation Plan for Revised Flexible Operating Hours Pilots

**28<sup>th</sup> June 2019**



## Contents

1	Pilot Rationale	3
2	Overview of evaluation approach	7
	Indicators relating to the <b>efficiency of use of court rooms</b>	9
	Indicators relating to <b>speed of justice</b>	11
	Indicators relating to <b>access to justice</b>	12
	Indicators relating to <b>working lives of legal professionals (solicitors, barristers, court staff and Judges)</b>	13
	Indicators relating to <b>experiences of public users</b>	15
	Indicators relating to <b>quality of justice</b>	15
	Indicators relating to <b>equality and diversity</b>	17
	Indicators relating to <b>costs</b>	18
4	Secondary Data Analysis	20
5	Bespoke data collection on utilisation of courts	21
6	Site visits to courts	22
7	Public user surveys	23
8	Qualitative depth interviews	30
	Interviews with professional users	30
	Interviews with support agencies	33
	Analysis of qualitative data	34
9	Cost Benefit Analysis, using a balanced score card	36
	Approach	36
	Step 1: Identifying the impact of flexible operating hours in pilot courts	36
	Step 2: Creating a balanced scorecard showing the impact of the pilots in the pilot courts	39
	Step 3: Identifying critical success factors	41
	Assessment of critical success factors	41
10	Reporting	43
11	Limitation of the evaluation	44
12	Risks and mitigation	45

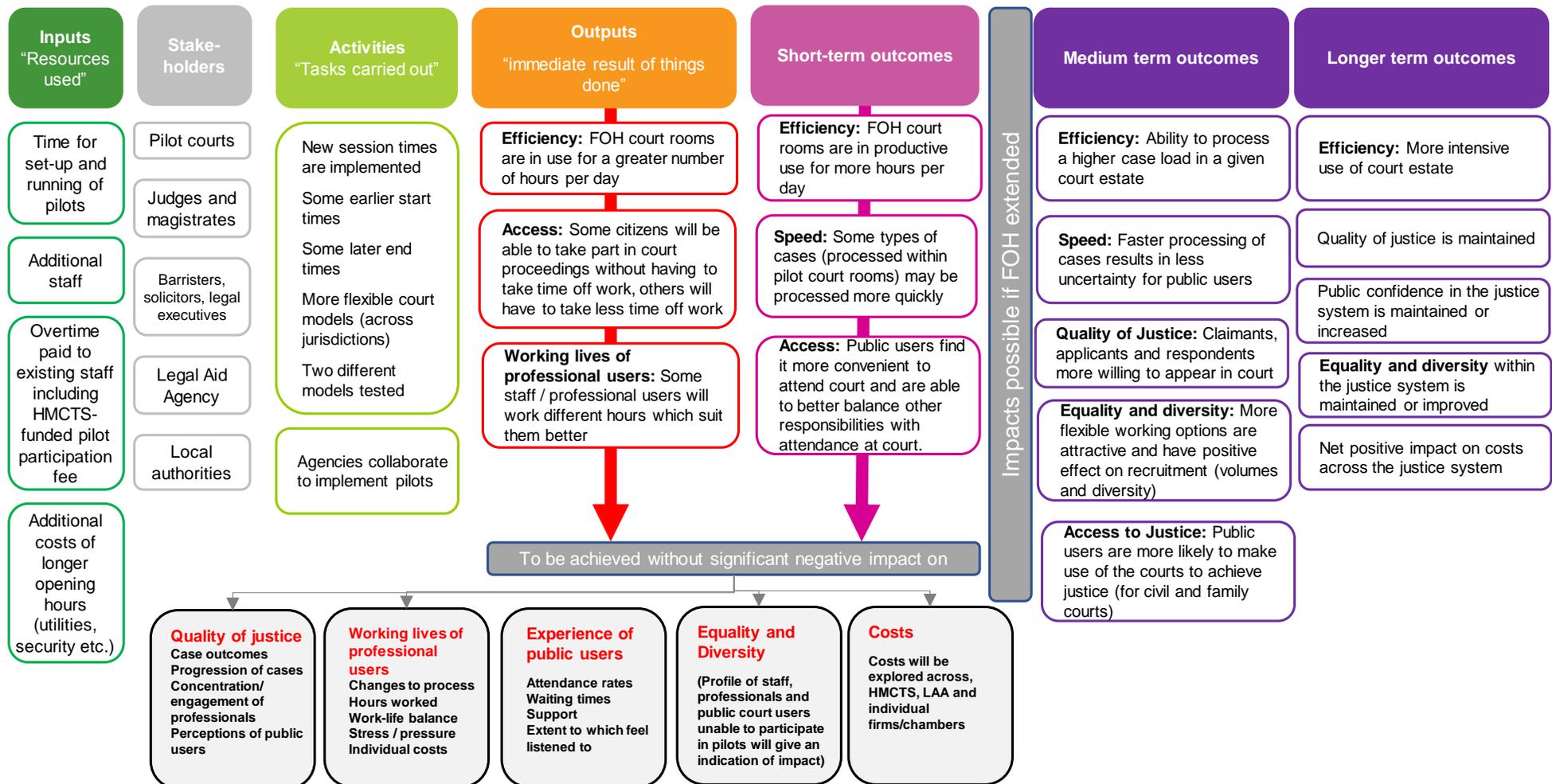
# 1 Pilot Rationale

- 1.1 Through a process of consultation with stakeholders we have developed a logic model for the pilot. This is a diagrammatic way of illustrating the outcomes and impacts that the pilot is looking to achieve. This model helps us to understand the key indicators that we will need to collect through the evaluation in order to draw conclusions about whether or not it has been successful.
- 1.2 The logic model is shown overleaf.
- 1.3 As the model shows, the primary benefit that the pilots are hoping to deliver is a **more efficient use of court rooms**. The intention is that longer operating hours will mean that a greater proportion of court time will be devoted to productive uses (to include hearings and box work), with less time where the court is not in use.
- 1.4 The other benefits that it is hoped the pilots might be able to deliver are:
- **Increased speed of justice.** A greater speed of processing could mean individuals secure justice more quickly.
  - **Improved access to justice.** Longer operating hours may improve the experience of public court users by offering greater flexibility (including less time off work). It may also mean that some users do not have to disclose attendance in court to their employers and might make it easier for some to balance responsibilities such as childcare with attendance in court. In the longer term this might encourage individuals to consider court action when they might otherwise not have done so.
  - **More flexibility for professional users.** Longer operating hours may be achieved without the need for longer working hours for professional users. Different working patterns might be established that may suit some better.<sup>1</sup>
- 1.5 The hope is, that it will be possible to deliver these benefits without impacting negatively on: quality of justice; the working lives of professionals involved in the pilot; the experience of members of the public when using the pilot courts; or the diversity of court users (discussed further in 1.6).

---

<sup>1</sup> Note that these potential benefits are not assumed to be true. They are hypotheses that we will seek to prove or disprove via the evidence we collect.

**Pilot rationale (“the problem to be addressed”):** The current usual sitting hours of 10am – 4.30pm (including 1 hour for lunch) do not make best use of the courts’ estate and are not convenient for public users.



1.6 At the same time, there is concern that the pilots could result in a number of potential disbenefits which include:

- **Additional costs** throughout the justice system;
  - **For courts:** Longer opening hours will increase the cost of running courts (more overtime, longer paid hours, more staff, utilities, etc). Longer opening hours may also increase the cost of the provision of support services.
  - **For solicitors, barristers and judges:** Longer opening hours may increase the cost of solicitors, barristers and judges required to service the courts (either more solicitors needed or additional need for overtime). Longer opening hours may also increase the cost to solicitors, barristers and judges required to service the courts (e.g. travel costs, childcare costs etc.) This additional cost may be passed back to the public purse.
  - **For the legal aid system:** In the immediate there is a possibility that a faster speed of processing could cause a one off “pull forward” effect on legal aid budget. Longer operating hours may increase the cost of legal aid payments. Longer operating hours may lead to Solicitors and Barristers needing to keep offices open longer and paying more overtime meaning they are likely to push for increases to legal aid payments (for the purposes of the pilot), a HMCTS-funded pilot participation fee will be paid.
  - **To public users:** Public users could incur additional travel costs if late or early starts change transport options that are available. They may also incur additional childcare costs.
- **A reduction in the quality of justice**, for example if there is greater strain on professionals affecting concentration levels, if greater levels of utilisation means there is less space available for client consultation or if earlier start/later finish times mean there is no time available for pre-hearing/post-hearing consultations.
- **A negative impact on the working lives of professionals** through affecting their work-life balance (which longer-term might lead to recruitment and retention problems) or impacting their working practices (e.g. managing preparation time and conferences out of court).
- **A negative impact on the experiences of public users** if early and/or later finishing times means individuals feel less safe travelling to and from court, greater utilisation rates means individuals feel less well listened to, individuals do not feel that they have had enough time with their representatives (as a result of pressure on space or time) or a lack of support services available for earlier/late starts mean individuals feel less well supported. There may also be difficulties posed by early starts/late finishes for public users with care responsibilities.
- **A reduction in the diversity of professional users.** It is possible that flexible operating hours will be less possible for / appealing to particular groups of professionals (within judiciary, solicitors, barristers, legal executives and court staff) and this in the longer-term could have a negative effect on the profile of professional users compared to the working age population. It is possible that flexible operating hours might lead to an unfair distribution of work – e.g. chambers allocating work in FOH courts to non-primary carers.

1.7 **In the evaluation it will be important to pick up any evidence of these disbenefits as well as of the intended positive outcomes.**

- 1.8 The logic model and evaluation approach set out in this document were developed by consulting a number of sources;<sup>1</sup> and with reference to discussions with, and representations from, a range of stakeholders, including via an Evaluation Advisory Group.<sup>2 3</sup>

---

<sup>1</sup> Sources consulted:

1. Initial pilot prospectus
2. Revised pilot prospectuses
3. Response analysis on pilot prospectus
4. Flexible Operating Hours Equalities Statement
5. Crown Double Shift Sittings Evaluation (Croydon pilot)
6. Lord Chancellor's Department, Extended Sitting Hours Pilot 2002 (Bow Street, London and Manchester; Magistrates' Court)
7. Nottingham Extended Sitting Days Pilot Evaluation
8. Process evaluation of the flexible criminal justice system pilots
9. Previous FOH pilots – summary reports with annexes (as provided with ITT for FOH Evaluation)
10. Bar Council Protocol for Court Sitting Hours
11. Brian Leveson's Review of Efficiency in Criminal Proceedings

<sup>2</sup> Stakeholders included representatives of the Sheffield Magistrates pilot, Brentford Civil pilot, , Manchester Civil and Family pilot; Legal Aid Agency, the Judiciary, Bar Standards Board, Family Law Bar Association, Resolution's Children Committee, The Law Society, The Bar Council, Bar Standards Board, DTUS, Shelter and CILEx.

<sup>3</sup> NB Over the longer term preparations to pilot FOH, criminal legal areas were included in plans, and therefore, relevant stakeholders were consulted. However, the criminal area of law is no longer included in the present FOH pilot plans. These former stakeholders' views and experience were considered through consultation in the formation of earlier plans for wider FOH evaluation and they are listed here: HM Prison and Probation Service; The Police; Prisoner Escort and Custody Service; Crown Prosecution Service; Blackfriar's Crown and Highbury Corner Magistrates Pilot; The Sheffield Magistrates Pilot.

## 2 Overview of evaluation approach

2.1 Taking into account the data that we need to collect to evaluate the pilots as well as the budget and timescales available we have arrived at a mixed-method evaluation design combining qualitative and quantitative research methods.

### Data sources

2.2 No existing data source provides us with the full range of evidence that is required to evaluate the pilots effectively. Our work to date has focused on understanding what we can likely draw from existing sources as well as developing an approach for collecting a range of other data from courts and wider stakeholders, where there are clear gaps.

2.3 We will have a range of different sources available for our analysis of impact. They include:

- Routinely collected MI or cost data;
- Routinely collected court data (not currently compiled centrally);
- Bespoke data collected from pilot courts;<sup>4</sup>
- Quantitative data collected from stakeholders for the purposes of the evaluation and
- Qualitative data collected from stakeholders.

2.4 Recognising that no data source is ideal, we have been forced to make difficult choices about where the biggest gaps in the existing data are and how we can fill them in a manner which is consistent with the scale, timeline and intention of the pilot. For example, data on utilisation of court rooms is a key area of interest for the pilots but there is no current data available for that assessment. We have therefore created a bespoke data template to collect that data from courts for a period preceding the pilot (a minimum of six weeks beforehand) and for the pilot period itself. As this data has never been collected before, it is not possible to be confident that a 6-week period is sufficient for a robust comparison between pilot and pre-pilot periods. However, based on our understanding of the flow of work in the two pilot courts on a monthly basis, which does not show great variation month by month, and an assumption that the courts will be able to collect accurate data over the periods involved, we are confident that 6 weeks is a reasonable period to develop indicative findings of impact. We will undertake cross checks using high level MI case data to ascertain whether there is anything particularly unusual about either the pre-pilot or pilot periods, such that the robustness of the comparison could be called into question.<sup>5</sup>

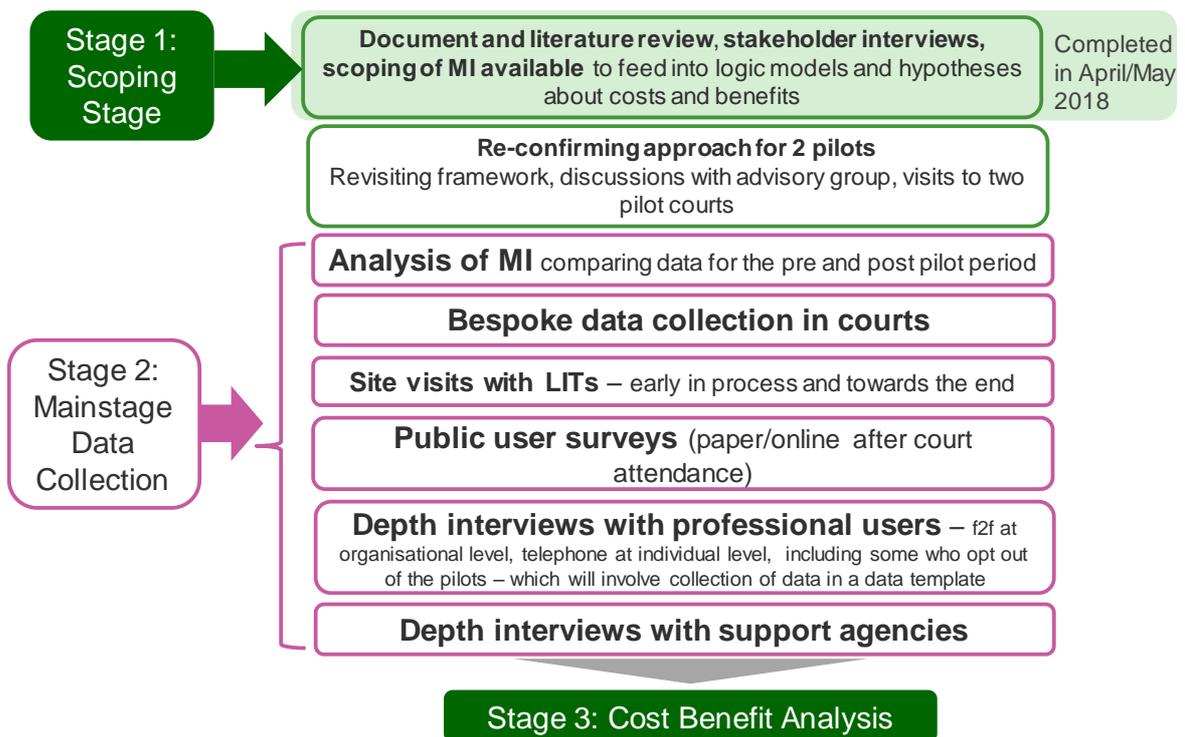
2.5 This also applies to our approach to the **counterfactual**, i.e. the data that we gather to assess what might have been expected to happen in the absence of the pilot. No one data source is ideal, and so we will need to blend routinely collected MI and cost data with bespoke data collected from the courts and from court users. This means that the data source used for the counterfactual varies according to what aspect of the pilot's impact we are examining. It also means that the counterfactual period varies too: some data sources (those already routinely collected) allow us to look at a longer period of time than others.

---

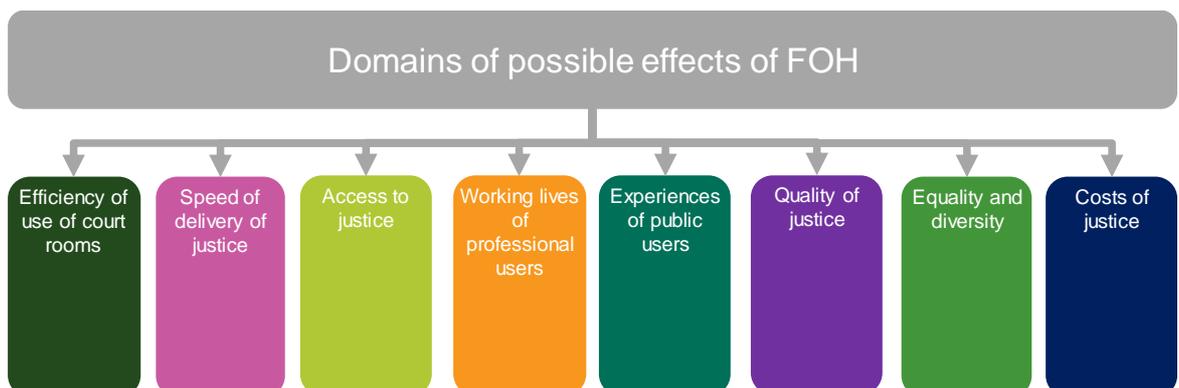
<sup>4</sup> This would involve one or more members of court staff on site, logging everything throughout the day in the pre and during-pilot periods.

<sup>5</sup> We could consider using post-pilot utilisation data as a contingency measure for our counterfactual, if there prove to be issues with the quality or amount of pre-pilot data collected. If pursued, this would add to analysis timescales and we would need to consider whether the working practices developed during FOH might have continued, post-pilot (for this reason post-pilot data are not ideal, and thus this is being thought of as a contingency only).

- 2.6 This brings us to the question of what sort of counterfactual group to use. In theory, it would be possible to identify, for each pilot court, a similar court that is operating ‘business as usual’ hours (using criteria such as profile and volume of cases, profile of professional and public users, transport links etc.) and designate this a match, for comparison purposes: having done this, our data could be compared between each pilot site and its match to assess what difference the flexible hours are making. We believe there is a key flaw in this however: a court that is a good match on paper may well be a poor basis for comparison for more subtle reasons (for example, the competence of key court staff members as managers or administrators). Our preference is therefore to instead compare data for the *same* court, prior to the introduction of flexible hours; with data for the period of the pilot.
- 2.7 The diagram below provides an overview of the different elements of the evaluation design, to generate the data needed.



- 2.8 Following on from the logic model, and our hypothesised benefits/disbenefits, there are a range of different **indicators** that we will need to collect for the evaluation, via the research elements above.
- 2.9 The indicators are grouped into the series of ‘domains’ shown below. These domains align with the labelling of outputs and outcomes in the logic model in the previous section.



- 2.10 The following tables show each of these indicators and the research element from the above design that we will use to collect them.
- 2.11 The tables include a column to show the counterfactual for each individual indicator. The counterfactual forms our point of comparison against which to judge the effect of the pilot. It is a measure of what might have been expected to happen to that outcome measure in the absence of the pilot. As our data sources are varied and have different strengths and weaknesses, we will not be able to use the same counterfactual for all the indicators of interest to this study.
- 2.12 When considering the counterfactual, it is important to note that our analysis will be performed at an aggregate level, not for individual cases. We will focus on whether there has been a change in the average outcomes observed between the pilot and counterfactual periods. For example, we will look at whether there has been a change in the average amount of productive court time in a given period (per day, week, 6 months) between the pilot and counterfactual period. For some indicators, it may also be appropriate to test whether a change has been observed in the range of outcomes observed or in the minimum or maximum outcome observed between the pilot and the counterfactual period. This will capture whether the pilot may have brought about a change in variability as opposed to a change in the average.
- 2.13 The tables below also include our view of the overall importance of each indicator to our ability to judge the success of the pilots. The intention is that our evaluation plan will cover all the listed indicators in some way but in the event that we have to prioritise (e.g. to stop interviews becoming too long) then we intend to prioritise indicators collected on the basis of this classification.

### Indicators relating to the efficiency of use of court rooms

- 2.14 Understanding efficiency is key to stakeholders. We will be looking to use a range of indicators for understanding efficiency:
- 1) Looking at the total number of sessions available and seeing how many of these have cases listed to them
  - 2) Looking at the total court room time available and seeing what proportion of that is used for listed hearings, what proportion is used for box work or unlisted hearings and what proportion of time the court is not in use (note that we are interested in the total productive time in which the court is used (hearings and box work), as well as the breakdown into different categories of use.
  - 3) By looking at whether there were any notable delays each day, and the causes of these (i.e. was it a 'pilot reason' or 'non-pilot reason')
  - 4) By looking at whether any cases listed for the FOH courtrooms were not heard, why that was (i.e. was it a 'pilot reason' or 'non-pilot reason') and what happened instead (e.g. the case was displaced to another courtroom, or was not heard at all)

Indicator	Source	Notes	Importance	Counterfactual
Total available/opening hours for pilot courtrooms	From pilot prospectus – checked with pilot courts		Most important	-

Indicator	Source	Notes	Importance	Counterfactual
Number of hearings listed into FoH sessions	MI collected by the LIT teams	This could be used to provide a measure of whether available sessions are actually used i.e. whether cases are listed into the FOH sessions.	Most important	Equivalent data from one or more 6 month periods comparable to the pilot.
Total volume and type of hearings in court	HMCTS CaseMan, FamilyMan and PCOL (case level data)	We will use data on this for all courts and for pilot courts in previous periods as a high-level <b>cross check</b> to identify broader trends across the system that we may want to account for in our analysis	Most important	Equivalent data for the pilot court from one or more 6 month periods comparable to the pilot. Plus, equivalent data for all courts for the pre pilot and pilot period.
Amount of “productive” time (time devoted to hearings or, box work) in the FOH court	Bespoke Data Collection in Courts	Bespoke template to collect information on the activity that the court is being used for throughout the day	Most important	Bespoke data collected in a baseline period (1-2 months pre-pilot)
Whether there were any notable delays, leading to increased time where court is out of work or not in use for hearings and reasons for these	Bespoke Data Collection in Courts	As above – and was it a ‘pilot reason’ or ‘non-pilot reason’	Most important	Bespoke data collected in a baseline period (1-2 months pre-pilot)
Whether any cases listed for the FOH courtrooms were not heard, reasons, and what happened instead	Bespoke Data Collection in Courts	As above – and was it a ‘pilot reason’ or ‘non-pilot reason’ For cases listed for the FOH courtrooms that were not heard, what happened instead (e.g. the case was displaced to another courtroom, or was not heard at all)	Most important	Bespoke data collected in a baseline period (1-2 months pre-pilot)
Number/proportion of cases starting on time by courtroom	Bespoke Data Collection in Courts	Possible that actual start and finish times could be collected for pilot cases (only listed times are available on HMCTS data)	Most important	Bespoke data collected in a baseline period (1-2 months pre-pilot) combined with information from CaseMan, PCOL and FamilyMan on listed start times

## Indicators relating to speed of justice

Indicator	Source	Notes	Importance	Counterfactual
Overall length of time to receive outcome	CaseMan, FamilyMan and PCOL data	Using information on “events” it is possible to identify the time to receive an outcome for a subset of all cases listed within pilot courts. By comparing equivalent case types before and after, we should be able to get an indicative view of whether there has been any effect on the length of time to an outcome	Most important	Equivalent data from one or more 6 month periods comparable to the pilot
Number of ineffective hearings <sup>6</sup>	CaseMan, FamilyMan and PCOL data on adjourned cases (CaseMan) and vacated casts (FamilyMan)  Plus Bespoke data template combined with listings data to assess reasons for listed hearings not going ahead	To establish whether this increases or decreases due to a pilot factor. We need to capture reasons (e.g. illness, non-service, parties not ready)	Important	MI data is limited so will provide a cross check Bespoke data collection for pre- and pilot periods (combined with listings data) will provide evidence source here
Number of profile days	HMCTS/ LITs data on number of profile days assigned to courts in pre and pilot periods  Qualitative interviews with courts and judges to investigate changes in profile days or ability to use profile days	Courts have a number of profile days available to them over the year. If the pilot period makes use of a greater number of profile days than expected for the period, this could have the effect of reducing the number of days available in other periods. <sup>7</sup> Further, some courts are not able to use all their profile days due to a lack of judicial resource so shifting profile days to pilot courts could potentially reduce	Important	Data from HMCTS/ LITs for pre-pilot and pilot periods  Qualitative views from courts and judges as to whether pilots have caused or are likely to cause impacts on available profile days or the ability of courts to use those days in any of the following (prior to the pilot

<sup>6</sup> As we understand this, HMCTs would define this as: on the hearing date, the hearing does not go ahead due to action or inaction by legal professionals or the court and a further listing for the case is required.

<sup>7</sup> We need to understand the number of profile days that the pilot court would normally receive and compare that to what they would have expected to receive either for the equivalent period (e.g. same period in previous year, previous 6 months, both). Pilot profile days/sitting days for FOH will be drawn on from the existing allocation for each site.

	compared to business as usual for court	available judicial resource for other periods or other courts. These could affect the speed of justice in other courts or during other periods.		period, during the pilot in non-pilot courts, after the pilots)
--	---	---	--	---

2.15 For indicators relating to **speed**, we will look at averages for each case type over the ‘pre’ and ‘during pilot’ periods in order to smooth out some of the case-by-case differences

### Indicators relating to **access to justice**

Data requirement	Source	Notes		Counterfactual
Time taken off work to attend court	Surveys of public court users		Most important	Survey of public court users – retrospectively sampling pre-pilot users
Perception of safety travelling to and from courts	Surveys of public court users		Most important	Survey of public court users – retrospectively sampling pre-pilot users
Availability of legal representation	Surveys of public court users	We should know from other sources whether a duty solicitor was available. Perhaps reasonable to ask individuals perceptions of availability as well?	Most important	Survey of public court users – retrospectively sampling pre-pilot users
Availability of legal advice	Surveys of public court users	We should know from other sources whether a duty solicitor was available. Perhaps reasonable to ask individuals perceptions of availability as well?	Most important	Survey of public court users – retrospectively sampling pre-pilot users
Perceived ease of travel to court	Surveys of public court users		Most important	Survey of public court users – retrospectively sampling pre-pilot users
Travel costs	Surveys of public court users		Most important	Survey of public court users – retrospectively sampling pre-pilot users
Caring arrangements that had to make	Surveys of public court users		Most important	Survey of public court users – retrospectively sampling pre-pilot users
Perceived convenience of time slot	Surveys of public court users		Most important	Survey of public court users – retrospectively sampling pre-pilot users

Data requirement	Source	Notes		Counterfactual
Time travelled to court	Surveys of public court users	Distance unlikely to change, length of journey may do	Most important	Survey of public court users – retrospectively sampling pre-pilot users
Time travelled to return home from court	Surveys of public court users	Distance unlikely to change, length of journey may do (due to different journey start times)	Most important	Survey of public court users – retrospectively sampling pre-pilot users
Number and proportion of public users opting out of FOH sessions (and selecting normal operating hours instead); reasons	MI collected by the LIT teams	We will need listing officers to record what cases ask not to be listed in extended hours and why	Important	Possibly no counterfactual available but may be possible to look at volumes of users applying to change hearing times.
Distance travelled to court	Surveys of public court users	Unlikely to change as result of pilots but useful to monitor.	Nice to have	Survey of public court users – retrospectively sampling pre-pilot users
Perceived convenience of location	Surveys of public court users		Nice to have	Survey of public court users – retrospectively sampling pre-pilot users

2.16 While surveys of court users (pre and during pilot) will be our principle source of data for indicators relating to access to justice, we will ask some questions of professionals (lawyers, HMCTS staff, judges) re: their perceptions of the impacts on public users

### Indicators relating to working lives of legal professionals (solicitors, barristers, court staff and Judges)

Data requirement	Source	Importance	Counterfactual
Proportion of staff working overtime	Site visits to courts Qualitative interviews with professional users	Most important	Records collected from professionals for both the pre-pilot and pilot period. We will signal the type of information required prior to beginning the pilot and provide a template to enable its collection.
Adjustments to workloads/responsibilities rearranged to accommodate need to work around FOH	Site visits to courts Qualitative interviews with professional users	Most important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Effect on hours worked (total, including preparation or boxwork time)	Site visits to courts Qualitative interviews with professional users	Most important	Interviewees' recall of the pre-pilot period (drawing on

Data requirement	Source	Importance	Counterfactual
			their own records, where relevant)
Effect on non-pilot work	Site visits to courts Qualitative interviews with professional users	Most important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Caring arrangements that had to make	Qualitative interviews with professional users	Most important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Perceived safety in travelling to/from court	Site visits to courts Qualitative interviews with professional users	Most important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Reasons for opting out of pilots	Qualitative interviews with those who have opted out	Most important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Arrival time prior to hearing	Site visits to courts Qualitative interviews with professional users	Important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Departure time after hearing	Site visits to courts Qualitative interviews with professional users	Important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Time to list out of hours hearings and deal with questions/applications	Site visits to courts Qualitative interviews with professional users	Important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Effect on support services who provide a presence in court	Site visits to courts Qualitative interviews with professional users	Important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Effect on agency/self-direction (i.e. ability to determine own time)	Site visits to courts Qualitative interviews with professional users	Important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Whether received sufficient food and drink and rest periods	Qualitative interviews with professional users	Important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
IT adjustments made to accommodate pilots	Site visits to courts	Important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)

### Indicators relating to experiences of public users

Indicator	Source	Notes		Counterfactual
Effect on trust and confidence in courts	Surveys of public court users		Most important	Survey of public court users – retrospectively sampling pre-pilot users
Waiting times	Surveys of public court users		Important	Survey of public court users – retrospectively sampling pre-pilot users
Number of litigants in person	Survey of public court users		Important	Survey of public court users – retrospectively sampling pre-pilot users
Whether litigants received food and drink, sufficient rest periods	Surveys of public court users		Important	Survey of public court users – retrospectively sampling pre-pilot users
Whether family/friends attended	Surveys of public court users		Important	Survey of public court users – retrospectively sampling pre-pilot users
Overall opinion of how well FOH worked for public users	Surveys of public court users		Important	Survey of public court users – retrospectively sampling pre-pilot users

### Indicators relating to quality of justice

Data requirement	Source	Notes		Counterfactual
Case outcomes	CaseMan, FamilyMan and PCOL data	Using information on “events” it is possible to identify the time to receive an outcome for a subset of all cases listed within pilot courts. By comparing equivalent case types before and after, we should be able to get an indicative view of whether there has	Most important	Equivalent data from one or more 6 month periods comparable to the pilot

Data requirement	Source	Notes		Counterfactual
		been any effect on the length of time to an outcome		
Failure to attend – who, how often	Combination of CaseMan, FamilyMan and PCOL data plus bespoke data collection	HMCTS data captures some information on why a hearing has not gone ahead; this is likely to be a partial picture so will be supplemented with data on reasons for hearings not happening/being delayed (in bespoke data collected by the courts)	Most important	Equivalent data from one or more 6 month periods comparable to the pilot
Use of duty solicitor (HPCDS)	LAA data on duty solicitor claims	Only relevant for a subset of hearings in pilot courts (HPCDS)	Most important	Equivalent data from one or more 6 month periods comparable to the pilot
Length of time to case resolution	CaseMan, FamilyMan and PCOL data	Using information on “events” it is possible to identify the time to receive an outcome for a subset of all cases listed within pilot courts. By comparing equivalent case types before and after, we should be able to get an indicative view of whether there has been any effect on the length of time to an outcome	Most important	Equivalent data from one or more 6 month periods comparable to the pilot
Professional users’ perceptions of impacts of FOH on quality of justice	Qualitative interviews with professional users	The interviews will explore how professional users define quality of justice, what they perceive the impacts to be and why	Most important	Interviewees’ recall of the pre-pilot period
Effect on concentration levels of professional users	Qualitative interviews with professional users		Most important	Interviewees’ recall of the pre-pilot period
Perceptions of quality of engagement with process (whether felt fully informed, whether had sufficient time with professional representatives, any indication of feeling rushed etc.)	Surveys of public court users		Most important	Survey of public court users – retrospectively sampling pre-pilot users

Data requirement	Source	Notes		Counterfactual
Number of matters listed and then vacated on application, and whether this was for reasons connected to the time of hearing	CaseMan, FamilyMan and PCOL data on adjourned cases (CaseMan) and vacated casts (FamilyMan)  Plus Bespoke data template combined with listings data to assess reasons for listed hearings not going ahead	To establish whether this increases or decreases due to a pilot factor. We need to capture reasons (e.g. illness, non-service, parties not ready)	Important	MI data is limited so will provide a cross check Bespoke data collection for pre-and pilot periods (combined with listings data) will provide evidence source here
Number of adjournments and reason for adjournment, and whether this was for reasons connected to the time of hearing	CaseMan, FamilyMan and PCOL data on adjourned cases (CaseMan) and vacated casts (FamilyMan)  Plus Bespoke data template combined with listings data to assess reasons for listed hearings not going ahead	To establish whether this increases or decreases due to a pilot factor. We need to capture reasons (e.g. illness, non-service, parties not ready)	Important	MI data is limited so will provide a cross check Bespoke data collection for pre-and pilot periods (combined with listings data) will provide evidence source here

### Indicators relating to equality and diversity

Indicator	Source	Notes		Counterfactual
Gender, age, ethnicity and disability status of all public users <sup>8</sup>	Public user survey		Important	Survey of public court users – retrospectively sampling pre-pilot users
Gender and age, profile of public users opting out of FOH sessions (and selecting normal operating hours instead)	Bespoke data collected by LITs		Important	Bespoke data collected in a baseline period (1-2 months pre-pilot)
Reasons given for deciding that a case in-scope for FOH was not suitable for an FOH session	Bespoke data collected by LITs	This relates to cases that the legal professionals decide are unsuitable. We would review to see	Important	Bespoke data collected in a baseline period (1-2 months pre-pilot)

<sup>8</sup> It has been mentioned that we are not covering all protected characteristics here. We need to be clear that this evaluation is not an Equality Impact Assessment and realistically it is not going to be possible to pick up changes relating to most protected characteristics – the sample sizes involved simply cannot be large enough to do so (to understand impacts on, for instance, perceived accessibility and convenience of Flexible Operating Hours sessions among protected characteristics groups, we would need sample sizes sufficient to support this analysis. This will mean focusing on groups who are likely to be more prevalent within the sample).

Indicator	Source	Notes		Counterfactual
		how many of these had a reason relating to equality and diversity		
Profile of court staff opting-in / opting-out of staffing FOH sessions compared to overall staff profile	Bespoke data collected by LITs		Important	Bespoke data collected in a baseline period (1-2 months pre-pilot)
Profile of judges opting-in / opting-out of staffing FOH sessions compared to overall staff profile	Bespoke data collected by LITs		Important	Bespoke data collected in a baseline period (1-2 months pre-pilot)
Profile of solicitors and barristers opting-in / opting-out of FOH sessions compared to overall staff profiles	Qualitative interviews with solicitors firms, barristers chambers,		Important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)
Instances of a change in advocate due to a FOH listing	Qualitative interviews with solicitors firms, barristers chambers,		Important	Interviewees' recall of the pre-pilot period (drawing on their own records, where relevant)

### Indicators relating to costs

Indicator	Source	Notes		Counterfactual
Change in costs to solicitors of operating in court (firm and personal costs)	Qualitative interviews with professional users		Most important	Records collected by professionals for the pre-pilot period. We will signal the type of information required prior to beginning the pilot and provide a template to enable its collection. This will include firm level cost information and personal costs such as childcare and travel.
Change in costs to Barristers of operating in court (firm and personal costs)	Qualitative interviews with professional users		Most important	Records collected by professionals for the pre-pilot period. We will signal the type of information required prior to beginning the pilot and provide a template to enable its collection. This will include firm level cost information and personal costs such as childcare and travel.

Indicator	Source	Notes		Counterfactual
Change in costs to Judges of operating in court (personal)	Qualitative interviews with professional users		Most important	We will capture (through interviews) any personal costs incurred by judges associated with the pilot courts (e.g. childcare and travel related)
Change in costs to HMCTS of operating courts	Input data will be collected directly from pilot courts (number and nature of staff required, other inputs) Unit cost of different types of court inputs available from HMCTS data Total cost will be imputed and compared to established HMCTS estimates	Total court cost data (e.g. utilities, staff) are paid centrally and then apportioned so not available on a per court basis. Estimates can be applied to pilot courts.	Most important	Court records on number of inputs (staff, other resources) required in pre pilot period. This will include payments for judicial time.
Change in LAA legal aid claims	LAA data on legal aid claims	There will likely be a time lag involved in this data due to claims processing times.	Most important	Equivalent data from one or more 6 month periods comparable to the pilot
Change in LAA claims related to duty solicitor scheme (HPCDS claims)	LAA data on duty solicitor claims		Most important	Equivalent data from one or more 6 month periods comparable to the pilot
LAA collated data on participation fee	LAA collated data on participation fee	Only collected during pilots	Most important	Counterfactual is zero payments
Change in cost of other organisations serving court (e.g. LA)	Qualitative interviews with professional users		Most important	Records collected by professionals for the pre-pilot period. We will signal the type of information required prior to beginning the pilot and provide a template to enable its collection.
Change in cost of court attendance for public users (to include changes in travel time/cost, time off work, childcare costs)	Survey of public users		Most important	Survey of public court users – retrospectively sampling pre-pilot users

2.17 The next few sections of this document discuss each of the elements of our overall approach, that we will use to collect and analyse the data for these indicators, in more detail.

### 3 Secondary Data Analysis

- 3.1 We will liaise closely with HMCTS and LAA colleagues to gain access to the relevant data for the pilot period and one or more appropriate counterfactual periods. For variables where the relevant indicator is well captured, the data should allow us to assess differences between the pilot and the counterfactual. For other indicators, the MI data will provide a broad cross check against the trends observed within the pilot.
- 3.2 We recognise that because some variables are only recorded in the MI sometime after the pilot hearing has occurred that this means that it will only be possible to conduct or confirm our analysis with a time lag following the pilots. We expect that time lag to be in the order of 1-2 months. This is true for data related to claims for participation fees, Housing Possession Court Duty Solicitor legal aid or wider legal aid incurred during the pilot period. We will work closely with HMCTS and LAA colleagues to ensure we are able to appropriately balance the need for representative data with the need for timely results.
- 3.3 The table below shows our preliminary view of what data is available and could be requested from HMCTS FamilyMan, CaseMan and PCOL databases for civil and family courts.

Category of date	Civil courts	Family courts
Case background	Case identifier Date of issue of proceedings Date of final order (if applicable) Court identifier	Case identifier Date of issue of proceedings Date of final order (if applicable) Court identifier
Hearing characteristics	Listed date Listed start time Session (am/pm) Hearing type Hearing outcome If not heard, reason	Listed date Listed start time Session (am/pm) Hearing type
Case type and complexity	Claim type (broad) Track (small claims, fast, multi) Claim amount	Case type (divorce/domestic violence, care)
Diversity and equality	-	Gender of applicants/respondents (some case types)

Source: Frontier drawing on HMCTS MI

## 4 Bespoke data collection on utilisation of courts

- 4.1 A key area of focus for the pilots is to understand whether the productive use of court rooms improves as a result of flexible operating hours. The HMCTS Management Information (MI) provides limited information for making this assessment.
- 4.2 Recognising this as a key gap, we have developed a bespoke data collection template to collect data on utilisation from pilot court rooms. This template will be rolled out for a period before the pilots (a minimum of six weeks beforehand) to collect a baseline measure of courtroom use as well as during the pilots. It will ask pilot courts to keep a daily log, recording the start and end time of all activity within the pilot court rooms (be this box work, listed hearings or unlisted hearings). Where the court is used for something other than listed hearings, the log will also ask for the reason for the use to be recorded. Furthermore, we will ask the logger to record any comments relating to what wasn't heard in the court, why this was, and what happened instead. The relevant judge will be asked to sign off on the accuracy of the daily log at the end of each session.
- 4.3 Combined with information from court listings, we should be in a position to identify the extent to which the pilot court was able to get through each of its listed hearings as well as what it was used for outside of these time periods. If unlisted hearings are included in the pilot court, we would expect to understand why that was the case. Moreover, if pilot court listed hearings have to be transferred to other court rooms, we would also expect to be able to capture this. Information about the wider utilisation of non-pilot courtrooms will be captured qualitatively in interviews with court staff and professional users.
- 4.4 In keeping with the scale of this project and the resource constraints of the courts, it is *unlikely* that we will be able to separately record any change in the utilisation of interview rooms or other parts of the courts as a result of the pilots. Any effects on these will instead be captured in the interviews and survey evidence.

## 5 Site visits to courts

5.1 To capture the views of those running the courts, we propose three site visits to each court to interview the Local Implementation Team (and anyone else on-site who plays a key role in implementation 'on the ground'). The time would be used flexibly, conducting a series of c.45-60 minute interviews, pairs, triads or mini-groups with the key individuals – determined by:

- Their availability;
- Whether individuals share responsibility for the same aspect(s) of delivery (interviewing these individuals together enables them to pool knowledge to create a complete picture);
- Whether individuals feel comfortable being candid in front of each other (for instance, we may wish to interview more senior staff separately to encourage candour among juniors).

5.2 At each court, we would seek to interview a mix of court staff involved in strategic decisions about the pilot; those involved in managing implementation; and those responsible 'on the ground'. A specialist IFF qualitative recruiter would contact and brief a senior manager at each court regarding the mix of individuals we seek (with this verbal briefing being confirmed in writing). They would then continue to liaise with each court until an appropriate programme of appointments has been agreed.

5.3 The three visits would comprise:

- One in the early stages of the pilot, to discuss the LIT's experience of setting up the pilot – how they envisaged it working; their rationale for decisions; any challenges; solutions to these; and lessons learned;
- One way mid through pilot operation to check on the reality of day-to-day running once the pilot is 'mature'; and whether any modifications to the overall approach have been made;
- A final visit at the pilot's conclusion, to explore their reflections on how well it has worked and why; and perceptions of the impacts (as well as any further changes made as the pilot progressed).

## 6 Public user surveys

- 6.1 Understanding the views of the ‘public’ court users is core to the evaluation and we believe that their experiences are best collected through surveys comparing the responses of those during the pilot period with those who have experienced as ‘business as usual’. While public users might, in qualitative interviews, be able to express views on how convenient or otherwise the new court hours were for them, what impact this had on their ability to access justice, and whether this might have been better or worse than the ‘business as usual’ hours, if this evidence comes from a small qualitative sample it will not be especially compelling. Furthermore, for public users, the occasion we are discussing will typically be their only experience of using the court in question, so they will be unable to compare how the experience of using the court has changed due to the introduction of flexible hours (since they will have no experience of ‘business as usual’ hours at the same court, with which to make comparisons). There may be positive or negative aspects of their court experience (sessions running smoothly, disruptions or delays), but they will find it difficult to attribute any of this to the operating hours.
- 6.2 Surveying ‘public’ court users quantitatively will not only provide more robust statistical evidence of the extent to which the flexible hours have made it convenient to access justice but also – by conducting surveys among a ‘control’ sample of court users (discussed below), we will be able to make statistical comparisons of public users’ perceptions of all aspects of the court experience, to answer the question, ‘with the flexible hours, was the user experience better, no different, or worse?’. Of course, for some indicators, ‘no significant change’ in user experiences between pilot and ‘business as usual’ hours would be a positive finding.
- 6.3 This brings us to the question of what sort of counterfactual group to use for the surveys. As discussed earlier, in theory, it would be possible to identify, for each pilot court, a similar court that is operating ‘business as usual’ hours (using criteria such as profile and volume of cases, profile of professional and public users, transport links etc.) and designate this a match, for comparison purposes: having done this, both survey data for court users and MI data could be compared between each trial site and its match to assess what difference the flexible hours are making. We believe there is a key flaw in this however: a court that is a good match on paper may well be a poor basis for comparison for more subtle reasons (for example, the competence of key court staff members as managers or administrators). Our preference is therefore to instead compare data for the same court, prior to the introduction of flexible hours; with data for the period of the pilot. This will involve us drawing a sample to survey court users both prior to and during the pilot. The key limitation of this approach is that it is vulnerable to any seasonal fluctuations in court experiences, but we believe it is the best available option.
- 6.4 A further consideration is how to access ‘public’ court users. We could either intercept users on-site and conduct interviews ‘there and then’, or draw samples to conduct interviews after the event. The advantages and drawbacks of each are shown overleaf:

Approach	Advantages	Drawbacks
<b>Site intercept interviews</b>	Recall of the court experience is fresh; once interviewers have clearance to access the site they can quickly access court users without long lead-in times to obtain sample.	Being interviewed may be an unwelcome distraction (or even a source of distress) while attending court; sampling will be less rigorous (as inevitably it will be driven by chance – court users who happen to be having downtime).
<b>Drawing a sample to conduct interviews later</b>	More ethical – the interview does not detract from the user’s business in court; a relatively complete sample can be drawn so that an attempted census can be conducted for the pilot period and there is a possibility of drawing a random sample for the pre-pilot period.	Likely to be relatively long lead-in times to obtain sample, accompanied by rigorous procedures to gain clearance to access and handle the sample.

6.5 On this basis, our recommendation is to conduct interviews after their court attendance.

6.6 The key groups of users to cover are:

- Claimants and respondents in civil cases (some of whom may be litigants in person)
- Applicants and respondents in family cases (some of whom may be litigants in person)

6.7 We anticipate that there will be a considerable lag between users attending court and their contact details being made available to us for conducting the survey due to the need for records to be compiled and data to be transferred. We have assumed a lag of around 2 months although we would be happy to work with a shorter lag time if it is possible to transfer records more quickly. To minimise the amount of process changes (other than the introduction of FOH) that might impact on the different experiences of users in the pre-pilot and pilot periods, we suggest drawing a sample of users from the period in the immediate run-up to the pilot launch for the pre-pilot survey i.e. those who attended court in March-May (we are estimating that we will need to draw our sample from across a period of at least 3 months in order to have a large enough starting sample) who we would then survey in July-September.

6.8 In order to obtain a large enough sample, we anticipate that we will need to draw our pilot users from across the whole pilot period (i.e. June-November 2019) and interview them in August-January.

6.9 We could sample our control group from a 6 month window before the start of the pilots to match the approach for the ‘during pilot’ interviews but this would make the pre-pilot period cover the Winter months whereas the March-May period feels a better match for the pilot period if we can generate sufficient sample records from this period.

	March	April	May	June	July	August	Sep	Oct	Nov	Dec	Jan
<b>Pre-pilot users</b>											
Attend court											
Interviewed											
<b>Pilot users</b>											
Attend court											
Interviewed											

6.10 Compiling the sample for this survey is likely to require accessing sample from a number of different places;

- Claimants and defendants in civil cases from the HMCTS CaseMan database
- Applicants and respondents in family cases from the HMCTS FamMan database

6.11 These are obviously very sensitive databases that need to be handled with care. IFF are very experienced in completing application processes to secure access to sensitive data. IFF also hold the ISO 27001 – the international standard for information security. All IFF staff have DBS clearance and some of the project team have enhanced Security Clearance through work conducted for the CPS.

6.12 The methodological options that are available for the survey depend on the contact details available. In 2015, the MoJ commissioned a team of independent researchers to undertake a study to explore the feasibility of conducting a representative and robust survey of civil court users<sup>9</sup>. The study looked at the suitability of using the HM Courts & Tribunals Service case management system data as a sampling frame and piloted telephone and postal survey approaches to test whether they could deliver robust and reliable data. As far as we are aware this study provides the best guidance for the most robust way to survey civil court users and the types of response rates that we might secure and hence we propose to mirror its recommendations about how to conduct the survey (and extend the logic to users of the family courts as well).

6.13 The key conclusions of this feasibility report were:

- That conducting a robust and representative survey of civil court claimants and respondents is challenging because of the lack of contact details on the sample frame;
- While address information was complete for the vast majority of users, there was limited inclusion of telephone numbers and e-mail addresses were rare;
- For the postal survey piloted, the impact of offering a £5 ‘thank you’ voucher was effective in increasing response rates to what the researchers felt was an acceptable level among claimants (a response rate of 31% was achieved among claimants when the voucher was offered compared to 22% without). For respondents the response rate was low at only 12% even with a voucher;

9

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/449575/civil-court-users-survey.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/449575/civil-court-users-survey.pdf)

- Based on these findings and an assessment of non-response bias, the researchers concluded that a postal survey of individual claimants would be broadly representative once weighted for non-response;
- The low response rate for defendants was considered too low for the findings to be considered representative;
- Response rates for a telephone approach were much lower (primarily because of a lack of telephone numbers on the sample file).

6.14 On this basis, we propose to conduct the survey work using a paper self-completion questionnaire mailed out to a sample of court users. In order to boost the response rate slightly and also to aid accessibility we suggest that the survey is also made available online. Alongside the paper self-completion survey, respondents would also receive an access code to enter their responses online if this is an approach that they prefer.

6.15 The feasibility study mentioned above experienced low response rates from respondents and this led them to conclude that it was not possible to conduct a robust survey with this group. Ultimately this means that if we include respondents within the survey then our achieved sample is unlikely to be representative of all respondents. There is then a discussion to be had about whether it is positionally better to collect at least some information from respondents or whether it would be preferable to have a more robust survey (i.e. one that could demonstrate an overall better response rate) which excluded this group. For the moment we have assumed that we will include them.

6.16 Our suggestion is that we conduct an attempted census of all users during the pilot period. The estimated numbers of users are shown in blue text in the table below. These volumes are based on calculations using the targeted case mix suggested in the pilot prospectus and the associated volumes of public users. **These initial calculations have been sense-checked by the Local Implementation Teams in the two pilot courts but still need to be verified by HMCTS analysts.**

6.17 Applying the response rates established in the feasibility study (for the approach using a £5 thank-you payment<sup>10</sup>) would generate the numbers of interviews shown in red text in the table.

Group	Likely response rate (if paper self-completion method)	Source of response rate estimate	During-pilot	
			Target no. of interviews	Estimated number of users
Claimants/ applicants	31% (with incentive)	Civil and private family court users feasibility study for MoJ	350	1111
Respondents	12% (with incentive)	Civil and private family court users feasibility study for MoJ	200	1601

6.18 For the pre-pilot periods we are likely to have a much greater volume of users to draw from (because we do not need to limit the sample to individuals whose cases were heard in the pilot court rooms as it seems reasonable to assume that the experience of those attending during the ‘business as usual’ period are a reasonable comparator regardless of the court room the case was heard in). In theory, we could therefore target a larger volume of interviews with users from the pre-pilot period, therefore reducing the confidence interval around these findings, which increases the ability of the study to detect change. The tables below show the impact of different sample sizes on the level of difference that we would need between the pre and during pilot surveys to be statistically significant at the 95% level.

Claimants			
Pilot users = 350 interviews			
Difference needed for statistical significance at the 95% level			
Size of pre-survey	For findings at 10/90%	For findings at 25/75%	For findings at 50%
350	5%	7%	8%
500	5%	7%	8%
800	5%	6%	7%
1000	5%	6%	7%

<sup>10</sup> Paying an incentive would be subject to MoJ/HMCTS Guidance and clearance. The feasibility study demonstrated that a ‘per survey respondent’ incentive was effective in boosting response, whereas there is no comparable evidence for the effectiveness of alternative incentives with this audience. Given the small populations involved, using a ‘per survey respondent’ incentive seems advisable in order to deliver samples that are as large as possible to maximise our ability to detect difference.

<b>Respondents</b>			
<b>Pilot users = 200 interviews</b>			
	Difference needed for statistical significance at the 95% level		
Size of pre-survey	For findings at 10/90%	For findings at 25/75%	For findings at 50%
200	7%	9%	10%
400	6%	8%	9%
500	6%	8%	8%
800	6%	7%	8%

<b>Combined</b>			
<b>Pilot users = 550 interviews</b>			
	Difference needed for statistical significance at the 95% level		
Size of pre-survey	For findings at 10/90%	For findings at 25/75%	For findings at 50%
550	4%	6%	6%
900	4%	5%	6%
1,300	4%	5%	5%
1,800	3%	5%	5%

6.19 There is no easy answer to the question of the volume of interviews with pre-pilot users that we target. Because this is a pilot exercise we can not be sure what level of change it might be reasonable to expect and hence what we would need the survey to be large enough to detect. Deciding on the final sample size will involve balance a range of factors including the overall budget for the evaluation.

6.20 The survey would need to comprise no more than 16 sides of A4 (and ideally we would aim for a shorter survey). We anticipate that it will need to cover;

- Waiting times
- Travel to court (time, cost, perceived safety)
- Whether time taken off work, how much and how this was accounted for (e.g. paid or not)
- Overall court experience
- Extent to which individual felt fully informed, whether had sufficient time/space to meet with professional representatives etc., any indications of feeling rushed etc.
- Time at which case was heard and overall perceptions of convenience of this
- Whether case started on time
- Preferences for times to attend court

6.21 The questionnaire would be drafted following sign-off of the overall approach. The questionnaire will be reviewed by the Evaluation Advisory Group (EAG) and MoJ, and revisions made prior to use.

6.22 For the quantitative survey of general public users, the key data outputs will be;

- Computer tabulations (with full significance testing against a key set of analysis sub-groups agreed in advance with the MoJ)
- An SPSS datafile

6.23 A full analysis of the data from each wave will be conducted, comparing results between those attending court in the pre-pilot and pilot periods. Analysis will be conducted at the overall level, for each jurisdiction and then for each pilot site and each user group individually.

## 7 Qualitative depth interviews

### Interviews with professional users

- 7.1 For the professional court users we are proposing a combination of telephone depth interviews and face-to-face depth interviews. We are proposing to use face-to-face depth interviews in cases where there are likely to be one or two key individuals to interview for each pilot (e.g. for local authorities, LAA, *firms* of solicitors etc.) from whom we will need to collect more complex information about costs and impacts **at an organisational level**. We propose to use semi-structured telephone in-depth interviews lasting 30-45 minutes for audiences where there will be a comparatively large number of respondents providing an individual perspective (judiciary, *individual* solicitors etc.) from whom we will be collecting information about impacts **at a personal level**. The rationale for a semi-structured approach for these audiences is to capture qualitative information while enabling coverage and analysis of the relatively large sample sizes needed to support analysis for each pilot site/model individually.
- 7.2 We still need to agree the best means of capturing details of professionals using the pilot courts. Our preference would be for records of attendance to be shared with IFF, from which to sample our interviews: if the details were incomplete (e.g. giving names without contact details), then we would use desk-based research to 'fill in the gaps'. We would ideally obtain records of attendance both before and after the introduction of flexible hours, to identify professionals with repeat experience of attendance, so they can make comparisons between the pilot and 'business as usual'. If obtaining lists of professionals attending court proves difficult, an alternative would be for us simply to approach local solicitors firms, barristers chambers, and screen for court users.
- 7.3 Our current thinking is to conduct 32 face-to-face organisational-level interviews and 29 telephone personal-level interviews structured as outlined below:

#### Face-to-face interviews

Site	The County Court at Brentford	Manchester Civil Justice Centre	TOTAL
Solicitors – as firms	10	10	20
Barristers chambers	6	6	12
<b>TOTAL PER SITE</b>	<b>16</b>	<b>16</b>	<b>32</b>

In addition, we propose to interview the Legal Aid Agency face-to-face:

	Number of interviews	Suggested distribution of these
Legal Aid Agency	2 pairs/triads	LAA staff involved in FOH across both sites / LITs
<b>TOTAL</b>	<b>2 pairs/triads</b>	

### Telephone interviews

Site	The County Court at Brentford	Manchester Civil Justice Centre	TOTAL
Judiciary	5	5	10
Barristers	7	7	14
Solicitors – as individuals	8	8	16
Duty Solicitor	3	3	6
CILEx Lawyer	3	3	6
Local Authorities	3	3	6
<b>TOTAL PER SITE</b>	<b>29</b>	<b>29</b>	<b>58</b>

7.4 Interviews will be a two-way dialogue in which the researcher responds to feedback from each participant, tailoring follow-up questions accordingly. They will be steered by a topic guide, to ensure all the key points are covered.

7.5 The topic guides would be drafted following sign-off of the overall approach. The topic guides will be reviewed by the Evaluation Advisory Group (EAG) and MoJ, and revisions made prior to use.

7.6 Our initial thinking about topic coverage for our ‘professional’ users is below. <sup>1</sup>

Topic	Broad question areas
Respondent background	<ul style="list-style-type: none"> <li>Their professional role; how this involves engaging with the pilot court</li> <li>What is their history of working within this court (over what period etc.)</li> </ul>
Initial opinion of pilot	<ul style="list-style-type: none"> <li>How was the concept of the pilot introduced to them (by whom, what were they told)</li> <li>What was their gut reaction to the idea and why (<b>for use as context to interpret their later answers re: how well pilot is running</b>)</li> </ul>
Overview of the impact of the pilot	<ul style="list-style-type: none"> <li>Now the pilot has ‘bedded in’, what concrete differences they have noticed in the running of the court, under the new flexible hours, compared with the usual hours</li> <li>Whether each of these is positive or negative and why</li> </ul>
Impacts in detail:	<p><b>Probing to explore the impacts on a range of specific issues – even if these are not spontaneously raised by the participant*:</b></p> <ul style="list-style-type: none"> <li>The effects on the smooth running of court sessions – e.g. number of adjournments, availability of necessary staff and participants, availability of rooms and functioning equipment</li> <li>The effects on the individual’s ability to do their job in court – what has been made easier, what has been made harder</li> <li>The peripheral / not-in-court impacts on them – such as travel to and from court, time to prepare, displaced workload, childcare needs</li> <li>The impacts on their personal ‘costs’ – is their time being used more or less efficiently and why (including whether waiting times have been reduced or increased); are their expenses increased/decreased</li> </ul>

<sup>1</sup> This is a high-level overview: the content of the full draft would be informed by the relevant metrics, discussed earlier (i.e. where the metrics needed are flagged as things to be captured via these in-depth interviews).

Topic	Broad question areas
	<ul style="list-style-type: none"> <li>• Their perceptions of the impacts on ‘public’ users – including effects on the ‘public’ experience of the court process; impacts on their ability to access justice; on quality of justice (e.g. how well-prepared professionals felt themselves to be); whether these impacts differ by type of individual (including groups with protected characteristics)</li> <li>• Their perceptions of the impacts on the type and volume of cases in ‘their’ pilot court</li> <li>• Their perceptions of the impacts on neighbouring courts, and reasons (e.g. potential changes in profile days)</li> <li>• Their perceptions of the impacts on partner agencies and other involved parties (including any international dimension)</li> </ul> <p><b>*specific issues to be tailored to the participant type</b></p>
Pilot management: early implementation	<ul style="list-style-type: none"> <li>• Thinking back to the early days of the pilot being introduced, whether there were any ‘teething troubles’ that have since been resolved</li> <li>• Whether aware of any adjustments made – and impacts of these</li> <li>• What lessons might be learned from this</li> </ul>
Pilot management: views on implementation when established	<ul style="list-style-type: none"> <li>• Now the pilot has ‘bedded in’, what is their opinion of the way the pilot implementation has been managed – and reasons</li> <li>• What has worked well – and why did this matter</li> <li>• What could be improved – and what difference would this have made</li> <li>• Any respects pilot has been tailored to ‘their’ court, and impacts</li> <li>• What are the key lessons learned from this</li> </ul>
Overall assessment	<ul style="list-style-type: none"> <li>• On balance, whether they consider the pilot hours positive, negative, or neither – and reasons</li> <li>• Has the pilot allowed the court to achieve anything it would have not have done otherwise – if so, what; whether this is significant or marginal as an outcome in their view</li> <li>• Whether there have been any unintended consequences – are these positive, negative, neutral; and reasons</li> <li>• Whether any modifications would make the pilots more successful in their view – explore idea(s) and rationale</li> <li>• What advice would they give MoJ regarding rolling this pilot out</li> <li>• <i>Interviewer recaps all of their lessons learned and suggestions: which would they prioritise, as being most likely to make a positive impact</i></li> </ul>
Final comments and wrap-up	<ul style="list-style-type: none"> <li>• Final comments from participant</li> <li>• Researcher gives reminder of any outstanding MI data being sought</li> <li>• Thank and close</li> </ul>

Using the professional depths to collect additional MI data

7.7 As discussed earlier, there is a challenge in that some of the MI data that we would ideally use to model the costs and benefits of FOH are not being collected as part of the work of the pilots (costs incurred by court users, including staff costs and expenses such as travel). For the ‘professional’ audiences that we are interviewing face-to-face, when arranging appointments, we will email out a datasheet of the information relating to costs that we would ideally collect during the interview. The datasheet will be devised by us and checked with the LITs during our early site visit (to ensure it is covering appropriate costs and ‘speaking the court user’s language’) and then signed off by MoJ.

Interviews with support agencies

7.8 A final strand of the evaluation fieldwork with ‘professional audiences’ is that of support organisations. We noted the ITT’s guidance that these may be covered in less detail, and therefore propose c.2 paired/triad interviews with these. Factoring in that these services applies only to certain of the pilot courts, our initial proposal is as follows:

	Number of interviews	Suggested distribution of these	Coverage by jurisdiction
Personal Support Unit	2 pairs/triads	Cover volunteers working with the public, and their co-ordinator, at the Manchester location	Civil Courts
<b>TOTAL</b>	<b>2 pairs/triads</b>		

7.9 Again, these discussions will be a two-way dialogue in which the researcher responds to feedback from each participant, tailoring follow-up questions accordingly. They will be steered by a topic guide – designed by IFF and signed-off by MoJ and with input from the EAG, to ensure all the key points are covered. Our initial thinking about topic coverage for the support agencies is below:

Topic	Broad question areas
Respondent background	<ul style="list-style-type: none"> <li>Their professional role; how this involves engaging with the pilot court</li> <li>What is their history of working with this court (over what period etc.)</li> </ul>
Initial opinion of pilot	<ul style="list-style-type: none"> <li>How was the concept of the pilot introduced to them (by whom, what were they told)</li> <li>What was their gut reaction to the idea and why <b>(for use as context to interpret their later answers re: how well pilot is running)</b></li> </ul>
Overview of the impact of the pilot	<ul style="list-style-type: none"> <li>Now the pilot has ‘bedded in’, what concrete differences they have noticed when providing services to this court, under the new flexible hours, compared with the usual hours</li> <li>Whether each of these is positive or negative and why</li> </ul>
Impacts in detail:	<p><b>Probing to explore the impacts on a range of specific issues – even if these are not spontaneously raised by the participant*:</b></p> <ul style="list-style-type: none"> <li>Availability of staff and volunteers for the session times</li> <li>Extent to which court sessions adhere to the listed times, and impacts of this</li> <li>Impacts on staff and volunteer travel to court (including costs and safety)</li> <li>Impacts on amount of time spent with public users, and quality of the time spent (e.g. whether in a conducive space, whether interrupted etc.)</li> <li>Other effects on staff and volunteers’ ability to do their job – what has been made easier, what has been made harder</li> <li>Staff and volunteer perceptions of wider impacts on public users’ court experiences</li> <li>Their perceptions of the impacts on neighbouring courts, and reasons</li> <li>Their perceptions of the impacts on partner agencies</li> </ul> <p><b>*specific issues to be tailored to the participant type</b></p>
Pilot management: early implementation	<ul style="list-style-type: none"> <li>Thinking back to the early days of the pilot being introduced, whether there were any ‘teething troubles’ that have since been resolved</li> <li>Whether aware of any adjustments made – and impacts of these</li> <li>What lessons might be learned from this</li> </ul>

Topic	Broad question areas
Pilot management: views on implementation when established	<ul style="list-style-type: none"> <li>Now the pilot has 'bedded in', what is their opinion of the way the pilot implementation has been managed – and reasons</li> <li>What has worked well – and why did this matter</li> <li>What could be improved – and what difference would this have made</li> <li>What are the key lessons learned from this</li> </ul>
Overall assessment	<ul style="list-style-type: none"> <li>On balance, whether they consider the pilot hours positive, negative, or neither – and reasons</li> <li>Whether any modifications would make the pilots more successful in their view – explore idea(s) and rationale</li> <li>What advice would they give MoJ regarding rolling this pilot out</li> <li><i>Interviewer recaps all of their lessons learned and suggestions: which would they prioritise, as being most likely to make a positive impact</i></li> </ul>
Final comments and wrap-up	<ul style="list-style-type: none"> <li>Final comments from participant</li> <li>Thank and close</li> </ul>

### Individuals who opt out of the pilots

- 7.1 To obtain a complete picture of pilot operation, we will also want to collect the opinions of professional users who opt out of pilot participation. This will be important, to capture these users' reasons for non-participation and (we suggest) explore what might have persuaded them to participate – thus identifying barriers and enablers that will be highly-relevant when considering pilot scalability. We suggest relatively small-scale coverage of those opting out: **10 in-depth interviews**, 5 per each of the site.
- 7.2 At the moment, we do not have full details about the different ways in which individuals might opt-out and what information will be recorded at the point when they opt-out. The approach for this element might need to be re-considered once we have full details (likely to be after initial conversations with the pilot courts).

### Analysis of qualitative data

- 7.3 The qualitative data will be analysed systematically using Excel frameworks. Our analytical approach involves the following steps:
- Interview write-ups**– using digital recordings, each interview will be written-up to a standard template by the interviewer conducting the site visit/interview.
  - Development of analytical framework** – this will be organised by key themes with data columns relating to each of the research questions. It is designed on the back of an analysis sessions with all those who have conducted interviews. It is designed to be organic so that it can be adapted as individual interviews are entered into it.
  - Entry of scripts into the analysis framework** – Interviews are entered into the framework by the research team. The first two interviews entered by each team member will be reviewed by the Research Manager to ensure consistency. The team will discuss any discrepancies and adapt the codes accordingly.
  - Review/analysis of the data.** Once all the data is entered into the framework, the information can be viewed horizontally (across individual interviews) or vertically within theme. Reports can be produced for different subgroups to provide insights into how different experiences and outcomes vary by independent factors, and to qualitatively

ascertain, for example, how views differ by professional group. This will also allow us to 'cut' the data by pilot variant (site/model) and by jurisdiction.

- 7.4 The analysis will be structured according to the logic model outlined at the beginning of this document, to address the key evaluation questions including whether the pilots were implemented as intended, and unpick the mechanisms by which the outcomes identified have been achieved.

## 8 Cost Benefit Analysis, using a balanced score card

- 8.1 The pilot is expected to have a multi-faceted impact spread across a wide range of indicators, affected organisations and people. The focus of this section is to set out how we will make use of the range of different data sources available, or collected specifically for this project, to assess the pilots.
- 8.2 The aim of the CBA is to quantify and value the impacts identified from the pilots across each of the domains of impact set out in section 2.9. It will consider both the magnitude of impacts observed and the distribution of costs and benefits across different parties (courts, agencies, individuals). We will also make an assessment of what can be inferred from the pilots about the possible impact of flexible operating hours in other courts across the country.
- 8.3 As it will not be possible to monetise or indeed prioritise amongst key metrics, we do not intend to present our CBA results as a traditional cost benefit ratio. Instead, we will use a balanced score card approach (where each indicator and domain of interest are treated equally), setting out the changes observed for each of the domains of impact, monetised where possible. The score card will show in what direction each indicator has changed and by how much, and who it affects.
- 8.4 If the pilot cost benefit paints a positive picture, where possible, our analysis will also seek to identify those factors that were critical to success in the pilots. This will inform a high level assessment of the extent to which the pilots are scalable and which factors are critical to the success of any future roll-out.

### Approach

- 8.5 We will take a three step approach to the CBA analysis:
1. Identifying the impact of flexible operating hours in the pilot courts (Brentford and Manchester)
  2. Creating a balanced score card showing the impact of the pilots in the pilot courts
  3. Identifying critical success factors to pilot success, if observed

We understand that it is very important that stakeholders recognise and accept the results of our analysis, particularly when we are describing the costs and benefits of flexible hours. We propose to involve stakeholders throughout the development of the CBA. We propose to use the Evaluation Advisory Group as the key channel for managing this engagement. We propose to share our assumptions and data with stakeholders transparently. We will ask them to validate and refine our assumptions, and to provide new evidence where they believe we can improve the analysis. In our experience this is a very good way to maximise buy-in from stakeholders, and also to generate additional evidence to inform our modelling.

### Step 1: Identifying the impact of flexible operating hours in pilot courts

- 8.6 In this step, we will bring together the wide range of different data sources that are available to understand the outcome of the pilots, as described in Chapter 2. For each indicator of interest, we will compare the observed outcome with an appropriate counterfactual to determine the impact of the pilot on that indicator. We will consider both changes in the average outcomes observed between the pilot and counterfactual period as well as changes in the range of outcomes observed.

- 8.7 A range of cross checks will also be performed to provide reassurance and account for, where relevant, any broader trends in the sector or seasonal effects.
- 8.8 The next subsections provide a high level summary of how we will consider the indicators described in Chapter 2 for each domain of interest. They are not intended to be a repeat of the detail captured in Chapter 2.

### Efficiency of use of court rooms

- 8.9 This data will be taken from a combination of the bespoke data collection template and MI collected by the LIT teams. The data will be used to ascertain the absolute and percentage changes in the total amount of productive time in court resulting from the pilots. This information will be accompanied by an assessment of the reasons behind any change such as a change in the number of cases starting on time, a change in the total hours available for hearings, a change in the productive use of existing court room hours etc.

### Speed of delivery of justice

- 8.10 We should be able to rely more heavily on HMCTS MI data (CaseMan, FamilyMan and PCOL) for indicators relating to speed of justice. We will seek to measure the following:
- the absolute or percentage change in the overall average length of time taken to receive an outcome on a case as a result of the pilots;
  - the absolute or percentage change in the number of ineffective hearings per case as a result of the pilots; and
  - the perceived change in the actual or potential future availability of judge time in non-pilot court rooms or courts or in pilot courts outside of the pilot (this could have the effect of slowing down delivery of justice outside of the pilot).
- 8.11 On the first indicator, we will compare the average length of time to outcome for cases pre and during the pilots to understand whether there has been any change. We recognise that for some longer cases, we may not be able to consider the length of time to outcome for all cases with one or more hearings in the pilot court. However, we will therefore consider, as our benchmark, only cases completed in an equivalent time period to the pilot.
- 8.12 On the second indicator, we are interested in whether changes to court opening hours that could mean that there is less time or space for pre-hearing consultations could result in a change in the number of ineffective hearings per case. On the other hand, the increased flexibility the opening times allow court users, could have the opposite effect.
- 8.13 On the third indicator, we are interested in whether the increase in profile days assigned to the pilot courtrooms could have the effect of reducing either the profile days or judge availability outside the pilot.

### Access to justice

- 8.14 Information on the indicators relating to access to justice will be captured from our survey of public court users during the pilots. Our counterfactual will be created from a survey of public court users for the period immediately preceding the pilot.
- 8.15 The indicators we will seek to measure are set out in detail in the table in Chapter 2. For each indicator we will seek to measure the absolute or percentage change in the indicator of interest.

For example, we will measure the absolute and percentage change in the number of hours taken off work by users to attend court.

### Working lives of professional users

8.16 For these indicators we will rely on a combination of site visits and qualitative interviews with professional users. Again, we will seek to measure all of the indicators set out in Chapter 2. For example, we will measure the absolute and percentage change in the proportion of court staff working overtime, the absolute and percentage change in the proportion of professional users working overtime etc.

### Experiences of public users

8.17 These indicators will be drawn from our public survey of court users. Again, we will measure the absolute and percentage change in each indicator set out in Chapter 2. For example, we will measure the absolute and percentage change in the number of public court users expressing their confidence and trust in the courts.

### Quality of justice

8.18 A combination of sources are needed to assess changes to quality of justice. We will again measure the absolute and percentage change in each of the quality of justice indicators set out in Chapter 2.

### Equality and diversity

8.19 We will rely on three sources for our assessment of impacts on equality and diversity:

- the interviews with professionals to understand whether there is any pattern in the profile of those opting out of the pilots or whether they have wider concerns;
- the public user survey to understand whether there has been any change in the mix of public users attending court over the pilot; and
- data collected by the courts recording the reasons why opt out decisions were made for cases, court staff, professionals or judges.

For some measures, such as the gender mix of public users, we will measure whether there has been a change in both the number and proportion of male court users and the same for female users. This will be repeated for age breakdowns, ethnicity breakdowns and by disability status.

For other measures, such as reasons given by court staff for opting out of FoH sessions, we will analyse whether there are any clear trends in the reasons presented (e.g. childcare responsibilities or travel concerns). We will not be able to present a change in these figures as there is no sensible counterfactual. However, these figures will be useful in considering the potential critical success factors for scaling up.

### Costs of justice

8.20 The costs associated with the pilots will need to be ascertained through a range of alternative sources, primarily:

- bespoke templates for courts asking them to set out the inputs they use (no. staff, hours of electricity etc) before and during the pilot (these can be costed up using average staff salaries, unit electricity prices for court available from MoJ cost data);

- information on claims for the participation fee during the pilot period;
- information on legal aid claims for Housing Possession Court Duty Schemes for pre and during pilot; and
- information on the costs incurred by solicitors, barristers and judges pre and during the pilot period, with a view to understanding any increment to those costs. This information will be collected during the interviews using a bespoke template. The template will look to capture the key inputs (e.g. staff time) required to operate in the pilot courts prior to the pilots as well as key changes in those inputs brought about by the pilots. The changes we expect to capture are likely to include changes to the number of staff required, number of staff hours worked or staff overtime paid and personal costs incurred (e.g. childcare, travel). We will cost these changes using both firm specific cost data as well as wider cost benchmarks.

For these measures, we will create metrics that capture the changes in cost for each relevant group (courts, professional users, public users etc) as well as at the aggregate level across all groups. We will be able to show the absolute and percentage change in costs associated with the pilots.

### Step 2: Creating a balanced scorecard showing the impact of the pilots in the pilot courts

8.21 This step of the work involves bringing together the evidence for all the metrics of interest across all domains of impact and reaching a balanced view as to the impact of the pilots.

8.22 Clearly, some of the metrics of interest can be monetised and others cannot. This is why we propose to use a balanced score card approach to represent the impact of the pilots. More on this below.

8.23 A number of the individual metrics outlined above can be combined into a single monetisable figure reflecting the government cost per hour of productive court time and how that has changed during the pilot. We have opted to use this figure within the CBA rather than trying to provide a financial measure of cost saving associated with any increase in productive time associated with the pilot. The reason for this choice is that the increase in productive court time is not a cashable cost saving in the short term. It simply translates into a capacity benefit until such a time as costs can be removed from the system.

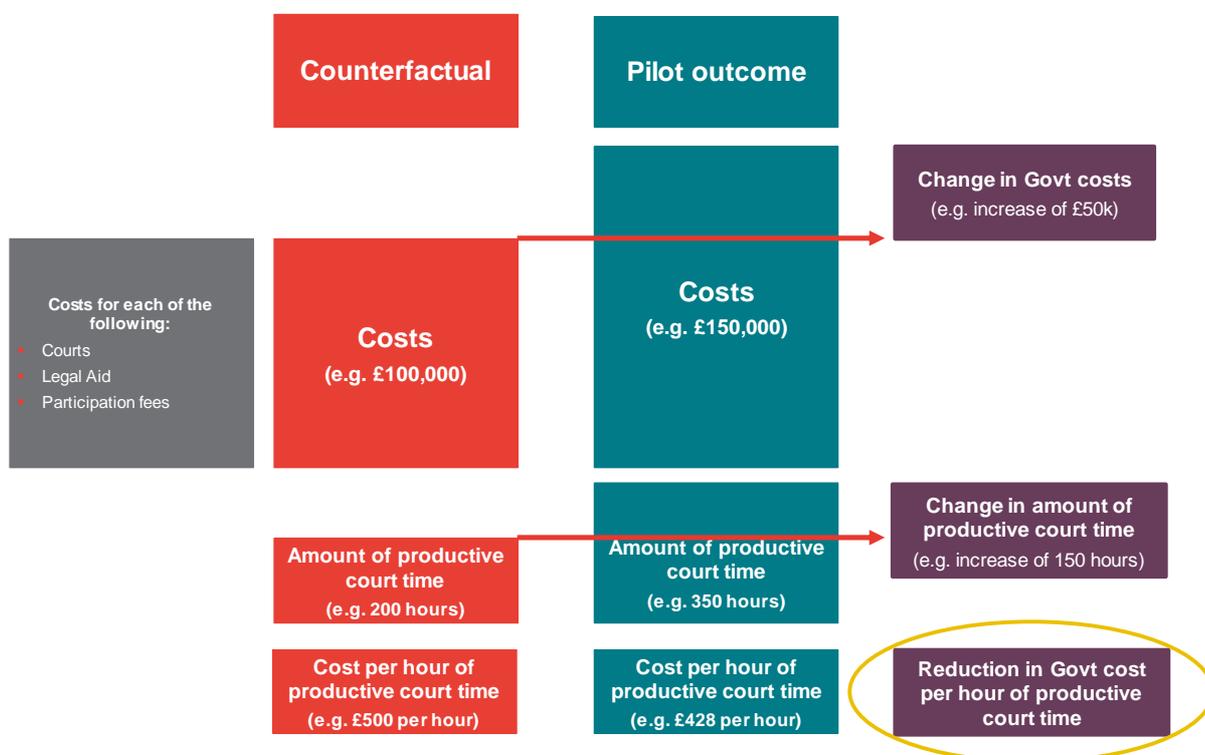
8.24 To calculate this metric involves bringing together information from two domains of interest:

- the change in government incurred costs associated with running the court over the pilot; and
- the change in the productive court time available over the pilot.

The figure below provides a stylised diagrammatic representation of this exercise (the numbers in the figure are for illustrative purposes only). The change in the cost per hour of productive court time resulting from the pilot is hypothesised to be negative in this example in line with the aims of the pilot.

To calculate this number, we will draw on the costs across all relevant categories both prior to and during the pilot period. We will also draw on the information regarding the number of hours of productive court time in the pre and during pilot periods. We can then calculate the cost per hour of productive court time pre pilot and the cost per hour of productive court time during the

pilot and compare the two to identify the change in the cost of a productive hour of court time as a result of the pilot. The cost per productive hour of court time will be calculated simply for those costs that are within the control of government (e.g. court costs, legal aid costs and participation fees). Wider costs such as those borne by professionals will be presented separately as part of the CBA.



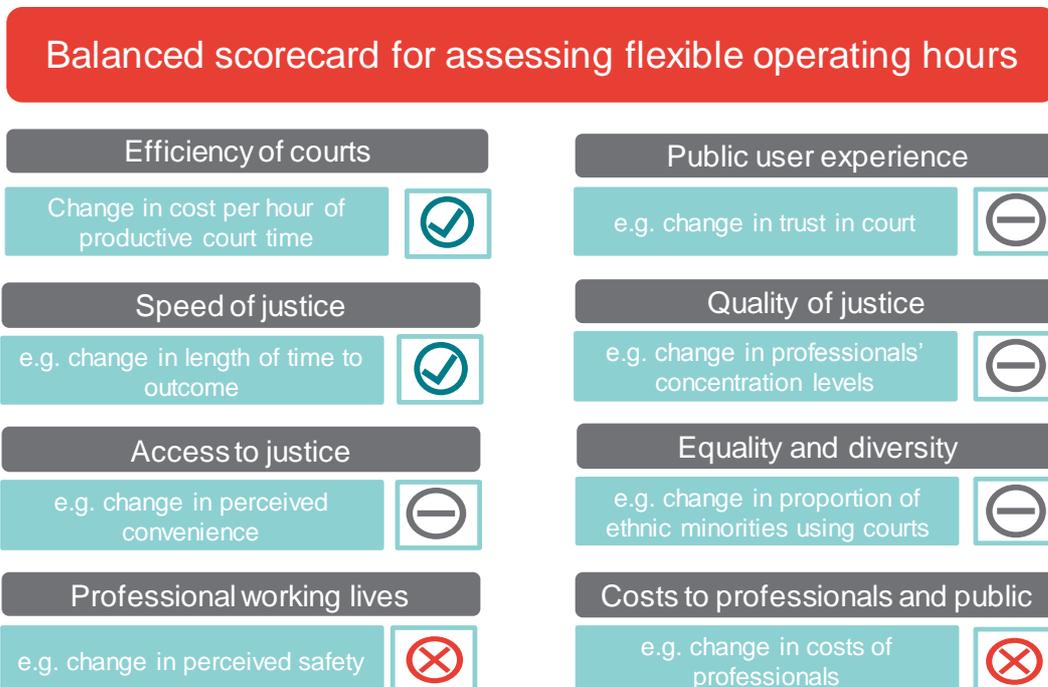
8.25 Many of the remaining metrics are much harder to monetise and in many cases this would be inappropriate. Therefore, we propose presenting our findings visually as illustrated below. The diagram below shows a simplified and stylised version of a balanced score card. As well as presenting the figures associated with each relevant indicator in each domain of interest (e.g. the change in perceived convenience of +X%), we will also indicate the relative sizes and directions of impacts within each domain. We treat each domain of impact as equivalent as there is no reason to prioritise any one domain of interest over another. For some domains, such as the efficiency of the courts, where indicators might vary for different groups, we will also highlight how the impacts vary across different groups involved in the pilot.

8.26 For those that are monetizable, we will use the standard approaches set out in the HMT Green Book to value those impacts. There are a number of important principles that will underpin any costs and benefits we are able to monetise.

- **We will not conflate system impacts with distributional impacts.** For example, total benefits might outweigh costs for a particular type of case but that might not be true for all parts of the court system (e.g. costs might be higher than benefits for legal aid agency even if total benefits exceed costs). We will set out the distributional analysis alongside the aggregate results to inform policy discussions.
- **Some costs are one-off (fixed) and others are recurrent (variable).** For example, the cost of communicating new hours may be a one-off cost, while keeping staff on duty for more total hours during the day may be a recurrent cost impact. Identifying the nature of these costs will inform the sustainability and scale-up potential of the pilots.
- **We will discount all costs and benefits as appropriate:** In addition, costs and benefits will be calculated over a set time-period into the future and discounted to Net Present Value (using

the appropriate Green Book discount rate), providing an indication of the possible future benefits from flexible hours. All of these technical details, and the best way to present them in the final results, will be discussed and agreed with the MoJ when we begin the CBA.

8.27 We will also provide an indication of the robustness of the evidence associated with each metric within the balanced scorecard approach so that conclusions can take account of the relative strength of each piece of evidence.



### Step 3: Identifying critical success factors

8.28 Once we have concluded the cost benefit analysis for the two pilot courts, we will assess what we have learnt about the potential for scaling up. As there are only two courts included within the pilot it is likely to be extremely difficult to give a clear sense of whether any potential roll out of the flexible operating hour model will translate to similar outcomes in other court locations. However, if one or both pilot sites do achieve a positive outcome from the pilots, we would make an assessment of the factors that have been critical to that success. Likely critical success factors might include the existence of the participation fee, timing of the flexible hour sessions, technology, size or geography of the area surrounding the court.

#### Assessment of critical success factors

8.29 Our assessment of critical success factors will draw on a range of evidence collected over the pilots.

- **Survey/interview questions:** Questions in the public user survey and qualitative interviews with professionals will ask about the extent to which specific pilot factors were critical to stated observations on impacts. We will also include a small number of scenario based questions in the public user survey and qualitative interviews asking about the likely direction and scale of changes to stated costs and benefits were a different set of factors to be applied e.g. a different set of timings, a different geography, a different set of cases or a different level of the participation fee.
- **Review of opt outs/design information:** We will review any evidence of opt outs from the pilot by professionals and public users to understand limitations to any roll out.

- **Review of variation in the model:** We will also make an assessment, where possible of any effect the different approaches used within the same pilot court over the course of the pilot or over the timetabled week made a difference to the observed outcomes.

8.30 An assessment of these factors will provide a starting point for consideration of how possible it would be to create the same set of factors in other locations. For example, we might be able to identify a subset of courts that might be best-suited to further flexible operating hours pilots.

## 9 Reporting

- 9.1 The core analytical outputs from the research will be a written report and an oral presentation of findings.
- 9.2 Our guiding principle in writing reports and presentations is to use the findings to tell a clear story. Within each section, our preferred approach is to structure the text with the 'headline' findings before providing the supporting detail. We believe the research findings should be accessible to a non-expert and presented to make the research topic as interesting as possible. We will deliver quality reporting – encompassing writing style, subtlety of analysis and quality of recommendations – primarily through the experience/seniority of the team writing/checking the outputs.
- 9.3 We will synthesize the findings from the various strands by theme, so that a particular topic is dealt with in its entirety, drawing on all of the relevant evidence, in a single section.
- 9.4 The debrief will be a PowerPoint document of c.30 slides. It will represent the findings visually, using text within diagrams, while always including 'headlines' to explain the main point being made by each slide, and highlighting 'take-outs' i.e. the practical implications for the end-user. The research directors would present.
- 9.5 The final report will be a Word document, still including visual elements but with greater emphasis on explaining the findings through text commentary. It will be more formal in its structure, with chapters and sections organised thematically. We envisage both the report and debrief containing:
- **Diagrams, to represent qualitative findings visually**, with liberal use being made of quotations and short case studies to place court users' 'voices' at the heart of the reporting output.
  - Plentiful use of **charts and tables for quantitative findings**. We choose a style of chart that best communicates the findings.
  - An **executive summary**, highlighting the key findings; plus **conclusions** and **recommendations for action** (we believe that it is our responsibility to push the 'thinking' about what the findings mean as far as possible, and so we aim to include actionable recommendations whenever possible).

## 10 Limitation of the evaluation

- 10.1 There are some limitations to what this evaluation will be able to tell us about the impact of Flexible Operating Hours in the courts.
- 10.2 The pilots are only operating in 2 courts and each individual court is different in terms of its physical structure, its staffing arrangements and its profile of professional and public users. Hence it is possible that impacts seen in these 2 courts might be different to those that would be experienced in others using the same model.
- 10.3 It is also worth noting that the pilot courtrooms are not just operating 'business as usual' for a longer time-period. The intention is to allocate very specific types of hearings into the extended hours slots. This means that an impact in terms of the profile of hearings and the speed with which particular cases progress has been designed in to the pilots. Hence the key positive impact that the evaluation will focus on is overall utilization rates. This approach to listing of cases into the extended hours slots means that findings could not be used to infer the potential impact of opening out extended hours sessions to other areas of law e.g. public family law work.
- 10.4 The pilots are also running only in 1-2 courtrooms within a larger court. In Manchester in particular these courtrooms represent a small proportion of the total number of courtrooms in the court. The FOH sessions are also only operating on some days of the week. This means that the additional hearing time generated by the FOH sessions is quite small at the level of the court and this may mean that it is difficult to pick up impacts at this level. However the FOH sessions add a relatively large proportion of additional hearing time at the level of the courtroom and hence it is at this level that we will be looking to assess impact.
- 10.5 One of the concerns about extending the hours of operation of the courts is that it will have a negative impact on the diversity of professional users. This will be very difficult to establish from the pilots (which will only run for 6 months) as these types of impacts tend to take years to show up. Within the context of the pilots there may be indications of an impact on equality and diversity (through looking at the profile of those that opt out for example) but it is unlikely to be able to establish definitively whether there would be an impact if FOH were rolled out more widely.
- 10.6 The pilot is expected to have a multi-faceted impact spread across a wide range of indicators, affected organisations and people. It will not be possible to monetise or indeed prioritise many of the key metrics which are likely to be affected. For this reason, we will not produce a traditional cost benefit ratio for the pilots. Instead, we will use a balanced score card approach (where each indicator and domain of interest are treated equally), setting out the changes observed for each of the domains of impact, monetised only where possible.
- 10.7 Overall these challenges mean that it is likely to be difficult to draw firm conclusions about the potential scalability of the pilots from this evaluation. However the evaluation should be able to determine whether or not there is likely to be value in exploring the FOH concept further.

# 11 Risks and mitigation

11.1 Our initial thoughts as to the main risks to the study and our mitigation approach to these are detailed in the table below

Risk	Impact	Likelihood	Action taken to reduce the risk	Action to be taken if the risk occurs
Lead times and administration involved in obtaining a sample of public users risks timetable	Medium	High	Putting a realistic, detailed work plan in place at inception meeting. Being available on a flexible basis (e.g. via ad hoc conference calls) to resolve any outstanding issues.	Review the timetable with MoJ, identify any flexibility in the timing. Given both waves of public user research are retrospective we could delay the fieldwork period by a week or so without jeopardising the overall reporting timetable.
Delays due to security clearances needing to be processed	High	Medium	This process was started as part of the original work (we looked to get it underway as soon as the project was commissioned) so we hope that we will be able to continue to use these clearances. We would welcome any assistance that the MoJ can provide with this process. We already have some staff on the team with SC clearance.	Basic disclosure only takes up to two weeks, so we would not foresee delays due to a need to get additional team members cleared to this level.
It is not possible to access lists of professional users	High	Medium	Our proposal assumes we can access attendance records to recruit professionals electronically – but we appreciate there is a risk that this might not be the case	We can look to compile a sample using desk research and approaching local firms, offices etc.
Not achieving desired volumes of public users interviews	High	Medium	Regular monitoring of case volumes during the pilot and available contact details. Ideally we would have a contact at the MoJ asap that we can talk to about securing access to contact details. Ensuring the survey is fully accessible – allowing people to complete online or by post if preferred.	Ultimately if the sample sizes achieved are smaller than ideal we should still be able to conduct the analysis – it is just that the difference we would need to see for a finding to be significant will be larger.
Pre and post waves of public users not comparable	High	Low	We'll conduct desk research to ensure there are no external factors which could introduce a bias and sense check this with professional users. Ensuring the pre and post pilot samples are drawn as close together as possible will also be vital.	We could explore the possibility of sampling public users from the same period a year ago instead (and accept there may be some recall risk).
Distress caused to public users participating	High	Medium	We have experience of working with potentially vulnerable audiences. Offering an online option to all respondents and ensuring they realise participation is voluntary. We are fully compliant with MRS and GSR ethical guidelines.	The team will have been trained on how to respond to distress. Signposting to appropriate support.
Insufficient capacity to conduct all the	High	Low	In total, we have allowed 8 weeks for recruitment and fieldwork of the professional audiences based on our experience this should be enough time to achieve all interviews. In terms of qualitative fieldwork	If necessary, we do have a further 10 qualitative specialists with relevant experience who could

Risk	Impact	Likelihood	Action taken to reduce the risk	Action to be taken if the risk occurs
qualitative interviews in the outlined timetable			capacity, we are confident that our nominated team can cover all the interviews in the required timeframe.	be seconded to this study (prior to MoJ approval) to ensure timescales are met.
The Hawthorne/observer effect	Medium	Medium	A risk that professional stakeholders feel strongly about the pilot (either positively or negatively) and try to influence the evaluation in a way that would not be replicated in a roll-out scenario. While we cannot necessarily reduce the risk, it is important to be aware of it.	We mitigate this through taking an approach to interviewing that is gently challenging (so for example asking respondents to back up claims about how well/badly the pilot has worked with evidence/examples of why that has been the case). This issue also highlights the need for quantitative data and factual questions to fully address the objectives.
MI doesn't exist for costs incurred by court users, including staff costs and expenses	High	High	This data will be central to the CBA and it is for this reason we have designed an approach where by the data is collected quantitatively from public users. For professional users, this data will be collected during qualitative interviews by asking users to fill in a pre-prepared template asking them about the resource involved in servicing the pilot courts and how that has changed during the pilot period.	To ensure accurate data is still collected from professional users we are proposing to warn them in advance, sending a data sheet so they can prepare ahead of our interview.
MI data provided by HMCTS has inadequate coverage or quality	Medium	Medium	Early engagement with Data and MI Group. Getting sight of this data as soon as we can will be important to determine an appropriate response. As most of the critical variables for this evaluation will be collected through bespoke mechanisms, the MI data largely serves as a cross check to ensure that the period of the pilot has not been affected by particular wider trends or seasonal effects.	Potential addition to primary research to gather targeted missing information. Where gaps persist we will estimate inputs as robustly as possible using other sources, and test validity with stakeholders.
Stakeholders do not provide sufficient data specific to their costs	High	Medium	Thorough planning and communication of primary research; work with HMCTS to agree most effective way to communicate and engage with stakeholders to maximise willingness to participate.	Estimate financial impacts based upon market assumptions, informed by wider sector evidence, stakeholder inputs, and Frontier Economics expertise in labour market characteristics.
Submitted outputs judged not to meet MOJ requirements	Medium	Low	We are very experienced at producing concise, well written and actionable reports and presentations for a very wide range of clients.	Should our draft outputs be judged unsatisfactory we will liaise with MoJ to fully understand where the problems lie and then produce new drafts which addresses these points and which is therefore in line with expectations.
Stakeholders reject findings	High	Medium	Stakeholders will be engaged from the outset, including regular meetings with the Evaluation Advisory Group. We will be as transparent as possible in explaining how findings have been arrived	Engage further to investigate and evaluate source(s) of concern. We would be happy to facilitate additional events to disseminate

Risk	Impact	Likelihood	Action taken to reduce the risk	Action to be taken if the risk occurs
			at. Our overall evaluation approach is also sensitive to differences between and within pilots, making it likely that findings will be nuanced rather than a blanket approval or rejection of FOH.	findings / discuss implications (at an additional cost). We would also be happy to advise on media strategy as necessary.
Data security breach	High	Low	IFF have stringent data security measures in place and we are ISO27001 accredited. This involves ensuring all sample files are held on secure servers with access restricted to the core research team. Furthermore, any transfer of data will be encrypted through PGP.	In the unlikely event of a data security breach we will immediately inform MoJ and commence an investigation to assess the extent of the breach and how it has happened. We would also suggest a face-to-face meeting with the MoJ team to discuss and agree steps to minimise the potential impact
Lack of cover for staff sickness, holidays and maternity/paternity leave	Medium	Low	The whole project team will be kept up to date of research progress. The impact of anyone leaving would be very short term. Having a 2nd director will ensure continuity at the head of the Project Team.	If necessary, we would replace members of the nominated project team with other similarly experienced IFF staff. We have the internal resources available to provide cover for all grades.

“

IFF Research illuminates the world for organisations businesses and individuals helping them to make better-informed decisions.”

Our Values:

### 1. Being human first:

Whether employer or employee, client or collaborator, we are all humans first and foremost. Recognising this essential humanity is central to how we conduct our business, and how we lead our lives. We respect and accommodate each individual's way of thinking, working and communicating, mindful of the fact that each has their own story and means of telling it.

### 2. Impartiality and independence:

IFF is a research-led organisation which believes in letting the evidence do the talking. We don't undertake projects with a preconception of what "the answer" is, and we don't hide from the truths that research reveals. We are independent, in the research we conduct, of political flavour or dogma. We are open-minded, imaginative and intellectually rigorous.

### 3. Making a difference:

At IFF, we want to make a difference to the clients we work with, and we work with clients who share our ambition for positive change. We expect all IFF staff to take personal responsibility for everything they do at work, which should always be the best they can deliver.



5th Floor  
St. Magnus House  
3 Lower Thames Street  
London  
EC3R 6HD  
Tel: +44(0)20 7250 3035  
Website: [iffresearch.com](http://iffresearch.com)

Contact details: Lorna Adams and Angus Tindle  
[lorna.adams@iffresearch.com](mailto:lorna.adams@iffresearch.com)  
[angus.tindle@iffresearch.com](mailto:angus.tindle@iffresearch.com)