Contents

1. Preface 1
2. Introduction 2
3. Start early 4
4. Define success 6
5. What to evaluate 8
6. Find a control group 10
7. Collect data 12
8. How long? 14
9. Plagiarise 16
10. Get everyone on board 18
Preface

In the spring of 2015, the What Works Centre for Local Economic Growth published a series of blogs on our website that introduced ways of improving the evaluation of local growth policy.

By that point, the Centre was half-way through our systematic reviews of existing evaluations, which sought to understand the causal effect of policy interventions and to establish their cost-effectiveness. One of the key messages emerging from our work was the need for better evidence on what works to foster local growth.

The blog series was intended to demystify the process of undertaking robust and useful evaluation of local growth programmes, and encourage more policymakers to undertake such evaluations. They were some of our most read posts in 2015.

This report collects the blogs as a single resource for policymakers to refer to when considering how to evaluate their local growth programmes.

For more resources and case studies on how to evaluate and evidence-based policy design, please visit the resources page on our website: http://www.whatworksgrowth.org/resources/

Henry Overman
Director, What Works Centre for Local Economic Growth
Introduction

By Meg Kaufman
Published on 27th March 2015

One of the greatest frustrations we have encountered in our reviews of the evidence on the impact of local growth policies has been the small number of robust evaluations from which to learn – particularly from the UK. As we suspected in our initial planning for the Centre, one of our most important tasks over the next two years will be to encourage more policymakers to undertake rigorous evaluation of their projects.

One reason people do not prioritise evaluating their projects is that they are not sure how to do it, and are concerned that it will be too expensive and time-consuming, drawing their effort away from programme delivery.

To help overcome this perception, we are promoting some simple steps that policymakers can take to dramatically improve the quality of their evaluations:

- **START early**
  - Evaluation is cheaper and more effective if it is integrated into project design

- **Define success**
  - Employment and productivity are good measures of local economic growth

- **What to evaluate?**
  - Do you want to focus on overall effects or on what works better

- **Find a control group**
  - Demonstrate your project’s impact by comparing participants with similar groups that have not been involved

- **Collect data**
  - Keep track of who participates and what happened to them before and after participating

- **How long?**
  - Short evaluations won’t capture all the effects, but longer ones are more expensive

- **Plagiarise!**
  - Use previous evaluations as templates

- **Get everyone onboard**
  - Give delivery partners clear instructions about what type of evaluation is acceptable
Another (less frequently voiced) concern about evaluation is that it may show that a project did not work as anticipated. Policymakers or their political bosses may prefer to use soft methods of evaluation that are not only easier to undertake, but more likely to show the project in the most flattering light.

This is an understandable response – no one wants to be associated with a failed project. However, it is important to everyone in the policy community that programme results, whether successful or not, are properly evaluated and the findings made available from which everyone can learn.

Stephen Curry wrote a piece in the Guardian recently about his decision to publish the results of a scientific experiment he recently completed, which had resulted in failure. He says there is often aversion to publishing stories of failure, but the desire to bury them must be overcome:

‘…negative results matter. Their value lies in mapping out blind alleys, warning other investigators not to waste their time or at least to tread carefully.’

This applies to economic development as much as it does to science, medicine or any other field. A policy that did not produce the results intended, or indeed failed, can provide useful guidance to others in the field if the evaluation is done well. Indeed, the opportunity to prevent others making the same mistakes can be the saving grace of a project that did not meet its original objectives.

More positively, when policy does have the desired positive effects, good impact evaluation helps convince others that the policy really is working: something that will be very important as local authorities seek to earn autonomy from a central government that is, rightly, sceptical about unsubstantiated claims. Experimentation and evaluation can also help improve policy effectiveness by focusing on policy design and increased understanding of what works better.

---

Start early

By Henry Overman
Published on 1st April 2015

In their focus on delivering a good policy on time and on budget, people often don’t think about evaluation until the end of a project (if at all). They then have to scramble around looking for ‘evidence’ to demonstrate its impact. Unfortunately, by this time it’s usually too late to provide convincing evidence of impact because the design of the scheme doesn’t allow for meaningful impact evaluation, or because basic data on the scheme (e.g. who got money) and how it operated (e.g. how decisions were made) are not available. In contrast, good evaluation is embedded in the policy design process. Sometimes this does happen in the UK,\(^2\) but not for most local economic growth policies.

In fact, a large portion of the evidence base we have found on local economic growth policies comprises studies which look retrospectively at projects that have finished. Many of these studies aren’t even official evaluations – instead being undertaken by academics long after the policy has been implemented (and possibly completed). Evaluations after-the-fact can be useful – as our evidence reviews demonstrate – but there are many benefits to be gained from considering how a project will be evaluated from the earliest phases of project design. For example:

Deciding how the scheme would be evaluated means thinking about what data will provide the most accurate picture of the impact of the project. In turn, thinking about what data would capture success (or failure) focuses minds on which objectives are the highest priorities. This can help prevent the tendency to attempt to address too many diffuse goals with one policy. For example, when thinking about evaluating an SME support programme we might think about collecting data on employment or firm productivity.\(^3\) Which is better would depend on whether the policy was trying first and foremost to improve profitability, or employment, or productivity. Asking what is the most robust method to evaluate progress against each of these objectives helps decide what is the main objective.

\(^2\) See for example some of the high quality evaluations discussed in the NAO report on evaluation in government, available here: https://www.nao.org.uk/report/evaluation-government/

\(^3\) See below on collecting data.
Thinking early about good evaluation approaches (such as RCTs or quasi-experimental approaches),\(^4\) can help design the programme and the evaluation in a way that provides more insight into how a policy works. In other words, this can allow us to evaluate what works better (i.e. how to improve effectiveness) rather than just focusing on what works (i.e. whether the programme had any effect). For example, in this Swiss employment support programme, vouchers for different monetary values were randomly assigned to businesses entering a programme. This design feature allowed the policy to be assessed not only in terms of overall impact, but also value for money. Policymakers were able to assess whether there were thresholds above which offering more money produced diminishing returns – knowledge which should allow for improved cost-effectiveness as the programme is developed. Importantly, everyone in this programme also received some assistance: clever design got round the common objection that randomisation leaves some people without help. Thinking about these issues carefully can also help improve policy processes – because it requires precision on exactly how decisions will be made about who gets what support.

Early clarity about how the project’s impact will be measured also helps with the evaluation itself. Firstly, it allows for the use of more robust impact evaluation methods, while taking in to account other constraints on the policy design process. Second, it helps to ensure that the proper data is collected throughout the project life. Arrangements may need to be put in place to track participants over time, or count visitor numbers, or measure changes in profitability. This data will be more cheaply and accurately collected if good evaluation is built into the programme design, rather than if it has to be reconstructed later. It also allows data requirements for evaluation to be considered alongside those for monitoring – reducing duplication of time and effort if the two requirements are considered separately.

Finally, embedding evaluation in the design process forces us to think about how the results of the evaluation will feed back into future decisions about the programme. In the extreme, this might involve decisions about whether to continue funding the programme. Alternatively, evidence on what works better (e.g. the appropriate level of subsidy) can help us improve cost-effectiveness as a scheme is developed. Evaluations long after the programme is finished don’t help with these two crucial decisions – and undermine the most important way in which evaluations may directly improve policy effectiveness for the specific policy being evaluated. Evidence reviews are useful, but certainly second best compared to a feedback loop direct from evaluation of a specific policy to the future development of that policy.

There are other advantages to embedding evaluation in the design process, but these are four of the big ones. The crucial message that emerges is to think about how and what you want to evaluate at the earliest stages of project development.

---

Define success

By Henry Overman

Published on 7th April 2015

The core objective of the What Works Centre for Local Economic Growth is to answer the question: which policies are most effective in supporting and increasing local economic growth. The answer, of course, depends on many things – not least what we actually mean by ‘increasing local economic growth’.

In looking at the evidence, and talking to local decision makers, it is very clear that there is plenty of disagreement about what is meant by increased local economic growth. In our evidence reviews, we tend to focus on two specific outcomes: employment and productivity. That is, we use a fairly traditional (some would say narrow) definition of local economic growth. We also worry about whether effects come about because of improvements for existing firms and residents, or because of changes in the composition of the local economy.

Quite rightly, some people worry that employment and productivity are not the main objectives of some policies that we review – e.g. estate renewal or the provision of sports and cultural facilities. But the reason for focusing our reviews on these two outcomes is precisely because we want to help local decision makers understand the most cost-effective way to increase local employment or productivity. After all, that’s our main remit.

All of that said, we also think it is very important to understand the wider effects of local policies on a whole range of outcomes. The crucial thing, especially when it comes to evaluating specific projects, is to carefully think through what effects we expect the project to have. To put it another way: in order to determine whether or not a project is a success, we need to think about what success would look like, and therefore how it should be measured. If, for example, estate renewal has objectives other than employment and productivity, then what are they and how would we know if we achieved them?

Unfortunately, even when the projects involved have a narrow economic focus, we consistently find that success is poorly defined. For example, in reviewing the evidence base for business advice effectiveness, we found a large number of assessments which were very muddled in stating the objective of a policy (if this was even defined). Policies often set out to support businesses to ‘grow’ – so far, so good – but when you start to measure business growth you have several choices. Are you interested in growing profits? Growing employment? Growing turnover? Output? Market share? Productivity?
Each of these factors is measured in a different way. More importantly, however, different forms of support will produce different kinds of ‘growth’. The importance of clarity on objectives is not merely an evaluation concern. If we want business advice to serve the end of local economic growth, we may need to look beyond the effects at the business level. Increased profits or market reach may not have much impact at all on the local economy if they don’t reflect increased productivity or translate into increased employment.

As we discussed in our post on ‘starting early’, one of the benefits of embedding evaluation into the design process is precisely that it forces us to think about how we might measure and define success. In turn, that forces us to think through exactly how we expect given policies to affect the local economy.
My last post dealt with the issue of the importance of defining success when thinking about how to evaluate. It might seem that answering that question also answers the question of ‘what to evaluate’. If the policy objective is employment, then we evaluate whether or not the policy has a positive impact on employment. If we are interested in cost-effectiveness, it is clearly helpful to undertake evaluations that address whether or not a policy works at all. It would be great if there were a large number of impact evaluations that clearly said that policy A works, policy B doesn’t. Unfortunately, as our reviews to date make clear (e.g. on employment training or business advice) this is rarely going to be the case – some employment training programmes work, others don’t (and similarly for business advice, access to finance, etc).

However, in practice, the challenge that we face in improving cost-effectiveness doesn’t just come down to conflicting findings. Call me cynical but even if the evaluation evidence on, say, business advice was uniformly negative, I’d be amazed if policymakers suddenly stopped funding programmes to provide such advice. Partly this comes down to politics – cost-effectiveness isn’t the only consideration when it comes to policy decisions. More prosaically, individual policy makers are often given a budget to spend on a broad policy area with one or two objectives. For example: commission an employment training programme to reduce unemployment among young people.

The question that such a practitioner really wants answered is not ‘what works’ but ‘what works better’? Given that we are going to provide employment training, what kind of training should we be providing? Short courses or long courses? On the job or off the job? Many considerations (resources, capacity constraints, etc) will have a bearing on how these questions get answered. In keeping with the Centre’s objectives, we would argue that cost-effectiveness should play a central role in answering questions about these policy design features. While often not well understood, good impact evaluation should play a crucial role in answering such questions.

For example, when NICE5 provides guidance it doesn’t try to answer the broad question about what makes us healthy. Instead, it tries to decide what treatments work best in addressing particular conditions. Similarly, the Educational Endowment Foundation6 focuses on assessing the effectiveness

5 https://www.nice.org.uk/
6 https://educationendowmentfoundation.org.uk/
of very specific interventions (including after school programmes, arts participation, extended school time, feedback) on improving one specific outcome: the attainment of disadvantaged pupils of primary and secondary school age.

Trialling two or more versions of a policy is a very effective method of comparing their effectiveness. Some of these experiments can be very large scale. One recent academic paper describes a trial from France involving a total of nearly 44,000 unemployed individuals allocated into three different groups. But trials can also be much smaller – which means that local policy experimentation can provide a great context in which to try to figure out what works better, particularly if a number of local areas are willing to collaborate in piloting different approaches.

We’ll be providing more detail on how to structure a policy’s implementation to provide comparison or control groups in the next blog.

---

How to evaluate - March 2016

Find a control group

By Henry Overman
Published on 23rd April 2015

Most policymakers are familiar with the basic approach to monitoring policy outputs (such as the number of people who have gone through a training programme, and how many got a job). I have done this hundreds of times in my career. However, in economic development we tend to be less practiced at isolating the causal impact of policy interventions.

By causal impact, the evaluation literature means the difference between the outcome for individuals ‘treated’ in a programme, and the outcome they might have experienced without it. Pinning down causality is a crucially important part of impact evaluation. After all, estimates of the benefits of a programme are of limited use to policy-makers unless those benefits can be attributed, with a reasonable degree of certainty, to that programme.

Establishing causality requires the construction of a valid counterfactual – i.e. what would have happened to programme ‘participants’ (individuals, firms or areas) had they not been treated under the programme. The way in which this counterfactual is constructed is the key element of impact evaluation design.

A standard approach is to create a counterfactual group of similar individuals, firms or areas not participating in the programme being evaluated. Changes in outcomes can then be compared between the ‘treatment group’ (those affected by the policy) and the ‘control group’ (those not affected by the policy).

Another approach – useful if we are interested in figuring out ‘what works better’ – is to offer similar groups different treatments. For example, in the case of business support, we take two similar types of business and offer some mentoring support (more expensive) and others are referred to online materials (a less expensive type of advice). The two groups can then be tracked to see how the different approaches play out.

The fundamental idea here is to make sure we are comparing groups that are similar. Such comparisons help address concerns that other factors might be driving changes for participants in a programme. These other factors might be ‘external’ to the programme (for example, it is offered through

---

to struggling areas) or they might be ‘internal’ (e.g. only certain types of firms choose to take part in a business advice programme). There is a lot more detail which can be considered when choosing a control group – and our selection of case studies for particular policies provide examples of how evaluations have dealt with these issues in practice. But even for those with less technical expertise – thinking about suitable comparison groups is a crucial step in thinking how you might evaluate policy success, as well as what you can learn in the absence of a comparison.

If you find yourself convinced, but are struggling to convince partners, you may find Ben Goldacre’s succinct description on Newsnight in February provides an invaluable introduction. And of course, we’re always available to try to help those looking to undertake more robust evaluation.

---

9 These are available on the Resources section of our site. A good example is this 2012 case study of a randomised control trial for employment training, available here: http://www.whatworksgrowth.org/resources/how-to-evaluate-employment-training-case-study/

10 Available here: http://www.badscience.net/2015/02/i-did-a-newsnight-thing-about-how-politics-needs-better-data/
Collect data

By Henry Overman
Published on 29th April 2015

It might seem simple, but the availability of appropriate data is often one of the biggest stumbling blocks for evaluating local growth policies, especially when trying to retrofit evaluation to programmes that are already well under way.

There are many points that could be discussed, but I am going to highlight three major ones.

First, you need information on the identity of programme participants. Which individuals, firms or areas have benefited from policy support? I think many people would be surprised that such information is often not readily available – even for accountability purposes if nothing else. But in the fifteen-or-so years I’ve been actively involved in thinking about the impact of government policy I’ve seen numerous occasions where such information is not being collected (even when large sums of money are involved). It can be a particular problem when second parties hand out the money on government’s behalf.

For what it is worth, my personal opinion is that such data should be gathered regardless of evaluation plans. But such data are critical to any attempt to understand the impact of policy. This information will need to be gathered and stored systematically and preferably in a way that allows it to be matched to other sources of information on participants (of which, more below). And if we truly want to get at questions of cost-effectiveness we also need to keep information on what kind of support participants have received (e.g. how much money) particularly if the kind of support can vary a lot across participants.

The second crucial data issue is to decide which data will best capture outcomes that are linked to the objectives of the programme. Thinking about how we define success should help clarify this. Once we’ve figured out the outcomes in which we are interested, there are two more things to think about:

- we need this data before and after the programme so we can see whether there has been any measurable change in the outcome for participants;
- we need similar before and after data for the control group that we are going to use as a comparison group for participants.

In fact, these two requirements – the use of before and after data, combined with a suitable control group – are key building blocks in our evidence reviews. Evaluations that don’t have both of these fall
short of the evidence standards we set for our evidence reviews as we discuss further in our scoring
guide.\textsuperscript{11}

The third decision that we need to make is how we will collect this data on outcomes. There are
essentially two possibilities. The first is to use bespoke survey data that is collected specifically for
evaluation purposes. Unfortunately such data can be expensive. It also brings a temptation to try to
collect large amounts of data – on the process, on a large range of outcomes, on whether people
think the programme is making a difference, etc. This then leads to large sprawling evaluation reports
which consider a huge range of issues. For most local decision makers it’s hard to believe that the
cost of bespoke data are justified.

I’d argue that we should focus attention on getting a smaller amount of data on key outcomes. At the
end of the day, in combination with a suitable control group, information on only one or two outcomes
such as changes in employment or wages allows us to answer the most important questions of all –
did the policy work or is one kind of support more cost effective than another?

Another option for keeping costs down is to make much better use of secondary data. Using such
data also helps address concerns about imposing a burden on the non-treated participants in the
control group as the data is already being collected for other reasons (small bespoke surveys also
help address that concern).

At the moment, using secondary data is easier for firms and areas than it is for individuals. For
example, for firms, if we have information on firm names, company house identifiers and postcodes
we can achieve very good matches with administrative data (such as the Inter-Departmental Business
Register) which already provides information on key outcomes of interest (particularly employment).
Even for individuals, efforts are being made to improve access to data sets – such as those held
by the Department for Work and Pensions – which provide detail on a range of outcomes. Indeed,
writing on our blog earlier this week, Majeed Neky described how this data – combined with a
randomised control trial – will hopefully allow a group of London councils to evaluate their ‘Working
Capital’ active labour market programme.\textsuperscript{12}

Of course, one problem with these secondary data sources is that there is usually a time lag before
they are available. Whether this is a major problem will depend on the time frame over which any
effects of the policy will be felt. For many local economic growth programmes, where effects are
expected to be longer term, this shouldn’t be such a problem – something we’ll discuss more in our
next blog.

\textsuperscript{11} Available here: http://www.whatworksgrowth.org/resources/scoring-guide
How to evaluate - March 2016

How long?

By Henry Overman
Published on 18th May 2015

In my last post in this How to Evaluate series I considered some of the issues around collecting data. One related issue that I didn’t discuss in that post is the question of when to evaluate?

In an ideal world, we’d want to think about when policy effects were likely to occur – and evaluate accordingly. For example, if the aim of a training programme is simply to get people into work it might be appropriate to evaluate shortly after the programme was completed. But if we’re interested in whether the programme leads to long-run employment, we might want to evaluate outcomes a year, or more, after programme completion.

In practice, unfortunately, there are often political pressures to provide early evidence on impact. It’s easy to criticise this impatience and to call for politicians to take a longer term view. But I think it’s important to try to move beyond this typical reaction – for two reasons. First, because experience suggests that appeals to take the long-term view will often fall on deaf ears. We don’t want the case for better evaluation to stand or fall on an argument we can’t win. Second, if we change our thinking to see evaluation as something that is embedded in the policy design process then the politicians may be right to insist on early evidence of impact (even if they are doing it for the wrong reasons).

In medical trials, for example, it is good practice for the trial protocol to describe the procedure regarding decisions on discontinuation of the trial. With that parallel in mind, and given the political realities, it is important to think about ways in which we could get some short-term indication of programme effects (possibly from bespoke survey data, or from secondary data that provides imperfect indicators) that will help inform policy development while waiting for secondary data to become available.

To give an example, let’s imagine that a LEP wants to provide information to Further Education students on the labour market outcomes they can expect if they take different types of courses. The idea being that telling them about, say, the high wages they will earn as an engineer might encourage them to take courses that lead to that career. The LEP decides to pursue a randomised control trial to evaluate this idea, before rolling out across all their colleges. To do this, they randomly send some new FE entrants information on likely labour market outcomes shortly before they choose their courses (there are issues with this very simple design, but we can ignore them for the purposes
of today’s post). The long-run aim of this policy is to increase the number of people employed as engineers. But there is plenty of opportunity for using short run indicators to see whether these long run effects are likely to be realised. FE colleges will have information on course choices, on completion rates, on exam outcomes and on first destinations. Using this information allows indicative evaluation within a few weeks (course choices), a few months (completion/results) and around a year (first destination for a one year course). If we see no movement on any of these outcomes, we might want to avoid wasting further money on this project next year (as well as avoiding the costs of the bespoke survey planned to pick up career destinations two years on from graduation). Of course, if these indicators are moving in the right direction, that may well justify continuation of the pilot and the costs of following up with a longer time scale.

The crucial point is that we need to be thinking about how the evaluation will be used to feedback into policy design. If piloting is truly embedded in the policy design process then we may want to make compromises on suitability to allow early evaluation based on readily available data. Of course, the long-run follow-up that looks at the ultimate objective is still vital to assessing the long-run effectiveness of the policy. But looking at short-run indicators can provide important information that can help determine whether to keep going – a decision which has implications for both programme and evaluation costs. Once again, this discussion only serves to highlight the importance of starting early and embedding evaluation into the policy design process.
One of the major problems with improving evaluation and getting it embedded in the policy design process arises because people think that evaluation is too complicated.

Throughout this series on how to evaluate, we’ve tried to show that thinking about evaluation need not be a very technical exercise. Asking simple questions about what and when to evaluate helps to define success. Thinking about how we’ll measure changes for those on the programme and compare those changes to a control group allows us to understand whether we will be able to attribute success to the programme (as opposed to all the other factors that will be at play). While thinking through these issues from first principles is to be strongly encouraged, we should also recognise the importance of copying freely from the approaches adopted in existing studies. Indeed, in an ideal world, we’d be able to draw on evidence from multiple randomised control trials before we even considered rolling out expensive interventions more widely.

Unfortunately, in practice, we are a long way from having that volume of high quality impact evaluation evidence available for any local economic growth policies. But our reviews do identify many evaluations that meet the Centre’s minimum standards – and thus provide a possible template for developing evaluation strategies. For example, this RCT shows how we might evaluate a scheme that provides adult education vouchers with the aim of improving labour market outcomes. A second example describes an RCT which uses business advice to achieve a similar outcome.

Admittedly, things get trickier when we move away from RCTs towards more statistical approaches. Indeed, one of the nicest things about RCTs as an evaluation approach is that they are relatively simple to understand, even for the non-specialist. Indeed, the EEF even has a simple guide for teachers to implement basic RCTs in their own schools and classrooms! But even when we move

13 Available here: http://www.whatworksgrowth.org/resources/how-to-evaluate-employment-training-case-study
away from RCTS there are lessons to be learned from existing evaluations - even if some of the statistical techniques are more tricky. Our web page provides a number of examples.\textsuperscript{16}

We also provide a methodology guide which allows organisations thinking about commissioning an evaluation to assess how different approaches might rank on the Maryland Scale that we use to rank studies for the purposes of our review.

In short – this is one area where plagiarism doesn’t necessarily give the right answer, but it certainly helps getting there.

\textsuperscript{16} These are available on our resources page: http://www.whatworksgrowth.org/resources/
Get everyone on board

By John Holden
Published on 10th June 2015

Previous posts in this series have looked at how good evaluation should be designed into local economic growth policies from the start and how, even with a small amount of planning, this needn’t be technically difficult or expensive. This post takes a look at a part of evaluation that is all too often overlooked: the need to bring the people or organisations you expect to act on the results in along with the evaluation process. Many good evaluations have failed to have the impact they rightly deserve because they missed out this crucial step. Likewise, many poor evaluations have had disproportionate impact because they got this right. There are two key questions that this blog explores: who needs to be “on board” and when.

The answer to the second question is easier than the first: as soon as possible, ideally before the intervention has begun or the evaluation design finalised. The reason to bring people in early is that by far the easiest way to ensure buy-in to the evaluation process, and ensure the results are acted on, is by developing a sense of ownership from the start. A few days delay in finalising the evaluation approach can save months of wrangling over technical details when the results of an evaluation are approaching publication. It can also make the difference between an evaluation being acted on and resulting in real policy change, and it being left sat on the proverbial (digital) dusty shelf.

When engaging with policy-makers early in the life of an evaluation, it is important to remember that most local economic growth policy making is short-term in the UK. Policy makers are unlikely to be interested in the design of an evaluation that isn’t going to report for perhaps two years or more as it doesn’t meet an immediate need. However, once publication is approaching, and the evaluation has the potential to disrupt an agreed policy approach, interest will rise exponentially. It’s best to pre-empt this and take the time to engage widely from the start, despite the indifference that may confront an enthusiastic researcher.

Deciding who to get on board is more complex. In local economic growth policy, the list of actors with a potential stake in an intervention is very broad. Rarely can one organisation act in isolation from its partners, as the interdependencies between different parts of the system are too great. The same is true of interventions and their evaluation. Engagement with four key groups of stakeholders is likely to be necessary:
• **The funding organisation:** When an intervention is externally funded it is likely that the funding organisation will require an evaluation and will in all likelihood fund it, so for purely pragmatic reasons it is helpful to have them on board with the approach. It is also the case that the funding organisation will usually be one of the key actors that the evaluation will be looking to influence, so their buy-in is essential.

• **The delivery organisation:** There’s no question that good evaluation requires access to high quality data and in most cases the delivery organisation will collect this information. Getting buy-in from those responsible for collecting evaluation data can make the difference between data input being seen as an unimportant administrative burden and it being viewed as an essential part of project delivery and service improvement. Buy-in from the frontline can also provide invaluable soft intelligence to give more rounded evaluation findings. A business advice intervention may fail because of a fundamental policy design flaw or because the business advisor spent all their time on Facebook. Spending the time building trust with those delivering a service can identify what the real lessons learnt are.

• **Partner organisations:** In local economic growth it is highly likely that any evaluation is going to identify that not all the benefits and not all the costs of an intervention flowed to the organisation that funded it. More sustainable delivery models can be developed if key partners are included in the evaluation process so that they can get an insiders understanding of the effect of an intervention on their organisations finances.

• **Key decision-makers:** In local economic growth, local authorities and local enterprise partnerships are likely to be important actors in deciding whether to scale up or scale down an intervention based on the evaluation findings. Early engagement can help ensure that they anticipate the evaluation findings and factor the timings of the results into their policy development cycles.

Finally, take some time to think about the person you will likely probably never meet but that is desperate to be on board with you: the researcher or policy-maker searching for evidence of what works. By making your evaluation findings available freely online you can have a policy impact that goes way beyond the immediate local context.
The What Works Centre for Local Economic Growth is a collaboration between the London School of Economics and Political Science (LSE), Centre for Cities and Arup.

www.whatworksgrowth.org
This work is published by the What Works Centre for Local Economic Growth, which is funded by a grant from the Economic and Social Research Council, the Department for Business, Innovation and Skills and the Department of Communities and Local Government. The support of the Funders is acknowledged. The views expressed are those of the Centre and do not represent the views of the Funders.

Every effort has been made to ensure the accuracy of the report, but no legal responsibility is accepted for any errors omissions or misleading statements.

The report includes reference to research and publications of third parties; the what works centre is not responsible for, and cannot guarantee the accuracy of, those third party materials or any related material.