

Are you sure?

Uncertainty in understanding impact, and where we might go next

Jess Dagers¹
Birkbeck College, University of London
May 2019

Impact measurement is essential for ensuring that resources flow to those businesses that best optimize impact and profit. If robust impact measurement standards are developed, we can prevent the moral hazard of ‘impact washing’: where businesses falsely claim to create positive impact, but in reality, achieve very little.

On Impact: A Guide to the Impact Revolution p13

These are the words of Ronald Cohen, in his manifesto for impact investing, “On Impact”.² The threat of ‘impact washing’ is a good place for us to start. It haunts the development of this new field, threatening to undermine the whole enterprise. The claim to be intentionally creating positive impact is what distinguishes impact investing from the mainstream, from the forms of investing that we have had for decades. Currently, the claim to be intentionally creating positive impact

is inseparable from measurement practices. So this is another paper about impact measurement, but I hope to convince the reader that it is bringing a fresh perspective to a debate that has been going on for some time.

‘Impact measurement’ refers to the information systems we build to give content to claims about impact. It is only by measuring, the argument goes, that we know what our impact is – otherwise we rely on anecdote, or good faith. Often the idea of impact *management* is also brought into play, to draw attention to the need to use the information about impact in running an organisation. In this paper our topic will be this emergent set of ideas around impact measurement and management (IMM) in the context of the attempt to build a global market for impact investing.

Of course, as soon as we are discussing ‘impact washing’ and the importance of IMM, it is not long before we hit on the matter of transparency and accountability.

¹Comments and feedback welcome. Please contact me at jessdagers@gmail.com.

²Ronald Cohen, *On Impact: A Guide to the Impact Revolution*, 2018

These terms draw attention to the need for all of us in the impact investing community to communicate the information gathered through IMM clearly and without bias, so that we can see what – if any – impact has been created. But transparency and accountability are also important because impact is about changing people’s lives, and those people who set out to change the lives of others should be accountable for the decisions they are making. There are plenty of people making these arguments, and the debate rumbles on over the best way to bring the enormous volumes of private capital to bear on persistent social issues.

This paper is based on my doctoral research. It is an attempt to communicate as succinctly as possible a series of observations and arguments about IMM, in the context of impact investing.³ I will argue that certain aspects of the attempt to produce knowledge about impact have been swept off the table, and that they need to be brought back to the centre of the debate if we are to have a chance of moving onto new ground.⁴

Before we get into the argument, I want to introduce a few brief snippets from the ongoing conversation around IMM that will be helpful later on.

NN Investment Partners

The first example is a statement put out by an impact investing fund called NN Investment Partners. As far as I am aware they are not particularly large or well-known players in the impact investing market. I came across this article by chance. I include this example because it demonstrates the kinds of statements that are now being made about impact. Under the title “Bigger and better impact universe”, the article states:

Recent proprietary research has re-

vealed that there are a surprisingly large number of stocks with a positive impact – almost 3,000 out of the 15,000 in our chosen universe (the CS HOLT database). ... Our results reveal that, contrary to conventional perception, positive impact companies tend to grow faster and have a lower cost of capital than those with a negative or insignificant positive impact.

*NN Investment Partners*⁵

Whenever I come across these kinds of claims about impact, I look for the method being used. In this case, the analysis proceeded as follows:

We first simply rated [the stocks] on their overall positive and negative impact. We then evaluated their positive and negative exposures to each of the individual 17 SDGs. Finally, we tagged the positive impact stocks to link them to our three themes (People, Planet and Prosperity) and eight underlying impact solutions. These company ratings are all judgement calls, and many are largely qualitative.

Ibid.

The reader might have their own opinions over the usefulness or defensibility of this kind of approach to understanding impact. I will not offer any opinions of my own at this point. I will just point out that this kind of analysis – where firms take time to classify their investments according to some kind of impact framework – is becoming a feature of the way impact investment is done, some of the time. I cannot make any claims about the prevalence of these kinds of practices, but it is uncontroversial to state that they are becoming increasingly common.

³This paper should be read in conjunction with a separate paper on the different logics contained in the idea of impact investing, forthcoming mid-2019.

⁴My argument is built on lengthy investigation of the connections between the logic of impact investing and the history of (Western) liberal, political and sociological thought. The ways in which we produce knowledge about phenomena (through practices such as IMM) are intimately bound up in the power structures shaping our lives. Further detail on the academic grounding of this argument is available in forthcoming publications and the full thesis ‘*Solving social problems with markets and measurement: a critical study of social impact investing*, (Daggers 2019), available on request.

⁵<https://www.nnip.com/en-INT/professional/insights/bigger-and-better-impact-universe> accessed 22nd March 2019

NPC and KL Felicitas Foundation

The second example is an impact report by a specialist impact measurement consultancy, UK-based New Philanthropy Capital (NPC), about one of the more prominent impact investors, the KL Felicitas Foundation (KLF).⁶ NPC were tasked by the Kleissners – the couple behind KLF – to draw together the impact data they had collected across their portfolio to build a picture of overall impact. We can expect this report to represent one of the better and more thorough examples of IMM, because of the prominence of both NPC and KLF. Notably, it presents the method as generally applicable – an approach to IMM that others can pick up and use themselves. It combines several frameworks in the analysis:

- the UNSDGs as a classification framework
- NPC's own 'Impact Risk Classification' (IRC) tool to 'compare impact practice' across portfolio companies, and
- the Impact Management Project's 'five dimensions of impact'

The report therefore does a number of things. It takes the impact data available from KLF's portfolio companies and attempts to aggregate the information against the SDGs. It uses the IRC tool to rate the impact practice across several dimensions, and it presents six case studies of portfolio companies framed within the IMP's fundamentals of impact. As is so often the case, NPC were faced with the task of consolidating impact data sets from multiple different settings. Their IRC tool is a way of tackling the incompatibilities they discovered – it is a scoring system that rewards commitment to "good impact practice".⁷ NPC suggests this scoring system can be used to understand not just how well an organisation does at measuring impact, but also the impact that organisation creates:

From NPC's experience of impact

measurement over the past fifteen years, we argue that a developed, intentional impact measurement process is likely to be associated with greater focus on impact, and by extension, an increased probability of impact.

In pursuit of deep impact, p18

In other words, NPC are suggesting that we can make comments about impact independently of any available data, focusing instead on whether organisations exhibit the kinds of practices and conventions that are associated with "greater focus on impact". The relevance of this example will become clear further down.

OECD

The third and final example is from a recent OECD report:⁸

The OECD SII principles recommend that data should be captured and aggregated in an internationally comparable manner as a means for further developing the evidence base for SII stakeholders and decision makers.

Social Impact Investment 2019: The Impact Imperative for Sustainable Development p216

The OECD here sets out an aspiration for the future of IMM: that we can develop systems that are coordinated and standardised, allowing aggregation and comparison across different settings. Importantly, the authors suggest this standardisation will emerge out of greater transparency. One of their proposed "transparency principles" is:

Transparency requires regular and standardised reporting by all who seek to be accountable for generating a positive impact from investments.

Ibid.

⁶Plum Lomax et al., *In Pursuit of Deep Impact and Market-Rate Returns: KL Felicitas Foundation's Journey* (2018).

⁷See page 20 of the report

⁸OECD, *Social Impact Investment 2019: The Impact Imperative for Sustainable Development* (2019).

Here we have comments on the same issues, from a different perspective: that of an organisation considering how to push forward the impact investing agenda.

We will pick up these examples again later in the paper. We turn now to the main argument.

Cause and effect

Here is our starting point: the idea of impact always contains three elements: an intervention (X), which causes a change from an undesirable set of affairs (Y) to an improved set of affairs (Z). The idea of creating impact is always about acting on the world to improve the lives of people, or the health of the planet.

This is not a controversial formulation. In fact it probably seems self-evident. But if we agree that impact can be characterised in this way, then there are implications for the kinds of knowledge we can produce about it, that is, for the way we do IMM. This claim will take a bit of unpacking.

We have stated that (X) is used to cause a change from (Y) to (Z). Understanding impact therefore involves two stages:

1. understanding whether (Z) has indeed been brought about. What outcomes have in fact been achieved?
2. understanding whether the change from (Y) to (Z) was caused by (X).

When it comes to impact, we want to know what causes the effects we observe, so that we know how to create the same effects again, and so that we know who should take credit for these positive changes (or, conversely, who should be held to account for negative changes).

Of course, this is the question of attribution, and I am certainly not the first person to raise it. Plenty of others have expressed opinions about how attribution should be dealt with in the context of IMM. We will come to these arguments in due course. At this stage I simply want to map out the contours of the problem. If we think about the issue of how we should produce knowledge about impact, I am suggesting, then the features

described above are always present.

To put it another way around, is it possible to conceive of impact without these three elements? Could we, for example, think of impact just in terms of observing a change from (Y) to (Z), without getting embroiled in the matter of connecting the change to (X)? I think it is quite clear this is a non-starter. Without connecting observed changes to the intervention in question, understanding 'impact' would be reduced to a purely descriptive account of things happening in the world, devoid of any sense of agency or control. Impact investing would be a matter of making investments and then reporting on a set of unrelated positive outcomes, without bothering to consider whether there was any connection between the two. It would be equally nonsensical to suggest a version of impact that did not reference a change of some kind. Impact has to involve something happening in a manner that can be differentiated from nothing happening at all. The question of cause and effect, therefore, is entirely unavoidably sat right in the centre of the attempt to generate knowledge about impact.

Introducing uncertainty

With the question of cause and effect comes uncertainty. There is always some degree of uncertainty over whether the change from (Y) to (Z) was caused by (X), or by some other factor. Scientific methods – to the extent we can generalise across them – attempt to reduce uncertainty as much as possible. Some of these methods get us close enough to certainty that we consider ourselves to have established 'the facts' about the effectiveness of an intervention, though philosophically (and statistically) speaking the uncertainty can never be

eradicated entirely.⁹

In the context of IMM, techniques such as randomised control trials (RCTs) are often pointed to as the most rigorous options for reducing uncertainty in results, and understanding attribution. This is one of the most contentious areas of the IMM debate. Some think the rigour of RCTs should be treated as a gold standard, a goal for anybody who truly wants to understand impact; others think RCTs are an expensive distraction that fail to generate meaningful or useful insight. But if uncertainty is inherent to our idea of impact, then it should always be a consideration, regardless of which method is being used to produce knowledge about impact. In this sense, the debate around RCTs and ‘academic rigour’ in impact measurement is something of a red herring. It suggests that the issue of uncertainty in IMM only applies when RCTs are being used, and that it doesn’t apply when RCTs are not being used. This is misguided. Instead, we need to recognise that uncertainty is present whatever we do.

If we accept this as a starting point, then we can get a clearer view on all the other considerations relevant when building an IMM system. As anybody knows who has tried to develop an IMM system themselves, there are many, many factors to consider alongside the demands of so-called ‘academic rigour’: how quickly are the results needed? Who will be using them, and what format do they require? How much time and money is available to develop the system? What kinds of data collection are feasible? How much detail is necessary? And so on. All of these choices have implications for how much certainty is possible. I want to be clear from the outset that I agree with prevailing wisdom in the sector that these considerations are important and valid. These restrictions are real, and cannot be ignored in pursuit of an unreachable ideal of rigorous research. But prioritising things other than reducing uncertainty does not mean that the uncertainty goes away.

My first point, therefore, is as follows: if we agree that uncertainty is inherent to the concept of impact, then we should also agree that uncertainty should be acknowledged whenever claims are being made about impact. This is a matter of transparency. Transparent communication of impact knowledge should mean faithfully communicating the degree of uncertainty in our results.

So what?

The reader might be wondering why this matters. Think back to the NN Investment Partners example described above. We only have a brief description of their method, but it seems fairly clear that they have not chosen an approach that seeks to reduce uncertainty. They explicitly use “judgement calls” to apply the ratings. And yet, they state their research “has revealed that there are a surprisingly large number of stocks with a positive impact”. Is this a statement that admits any uncertainty in the finding that the stocks have positive impact? No. It is stated as fact.

I think this is problematic. NN Investment Partners have chosen a method that generates knowledge containing huge amounts of uncertainty, and yet they make statements of fact. The point here is not that they should have used a different method, but that they have not recognised the limitations of the method they have chosen explicitly enough.

NPC’s report faces similar challenges. One section of the report maps the KLF portfolio against the SDGs. They present a table containing a series of figures which “gives an overview of the impact of the KLF portfolio in each area”. Under SDG 1 “No poverty”, for example, they report “37.3 million people with access to basic services” and “over 26 million people benefitting from cost savings generated” (p26).

I also find this problematic. Even though, two pages later, they clearly acknowledge that they “have not

⁹For an introduction to the philosophical conundrum surrounding cause and effect, see Judea Pearl, ‘The Art and Science of Cause and Effect’ (1996) <http://bayes.cs.ucla.edu/BOOK-2K/causality2-epilogue.pdf>

made any assessment here about the robustness of the data or assessed how reported data is linked to mission or impact” (p28), the presentation of the information makes it possible to read these metrics as ‘KLF’s impact’ without paying any attention to the enormous amounts of uncertainty they contain.

This problem is compounded by the Impact Risk Classification tool. The IRC could quite feasibly be used as a system for assessing the state of impact measurement practice within an organisation. But NPC, when they say “a developed, intentional impact measurement process is likely to be associated with greater focus on impact”, suggest that scores allocated within the IRC also tell us about the impact that organisation creates. On the contrary: relying on this assumption introduces so much uncertainty into the analysis that the IRC is unable to provide any meaningful insight into the impact being created. Transparency over uncertainty would require that these limitations are clearly set out.

Fit for purpose?

Here is the next stage of the argument: what counts as good quality IMM depends on the purpose for which IMM is being done. Again, this might sound obvious, but the implications are significant.

Importantly, this stage of the argument relies on the previous discussion. We must be very clear in separating out the questions of causation and uncertainty from the question of what counts as ‘good’ or ‘bad’ IMM. In other words, ‘good’ IMM does not have to mean IMM that reduces uncertainty using scientific methods. Similarly, the most robust, certain results can still be seen as ‘bad’ IMM. What matters is whether the IMM system is fit for purpose. In turn, what makes an IMM system fit for purpose depends on the particular setting in which it is being developed.

I need to introduce some new terminology to properly communicate this stage of the argument. We have already talked about IMM as the attempt to ‘produce knowledge’ about impact. In my academic writings I

Is this anything more than semantics? This paper started with Ronald Cohen’s description of impact washing: “where businesses falsely claim to create positive impact, but in reality, achieve very little.” On a first reading, this suggests we should be concerned about businesses being disingenuous, knowingly making misleading claims to impact. I imagine this is what Cohen is referring to. We can now add a second, far more worrying interpretation of impact washing: where businesses with the best of intentions genuinely think they are creating a positive impact when in fact they are not. If IMM continues to admit huge amounts of uncertainty in statements about impact – uncertainty that is unacknowledged, and not even understood by those producing this knowledge – then the risk of impact washing is exceptionally high.

We will return to this issue below, but first I want to turn to a second set of basic features of the attempt to produce knowledge about impact.

use the term ‘orientations to knowledge production’ to capture the sense in which what counts as good quality knowledge varies from one setting to another. That is, the processes used to create knowledge can be oriented to a variety of ends, and this affects what counts as good or bad quality.

We can use examples from outside impact investing to illustrate this point: cost-benefit analysis (CBA) was developed as a way of supporting decision-making about public works, estimating the costs and benefits of different options for (say) the location of a new bridge. In contrast, survey sampling techniques were developed for increasing the accuracy of our knowledge about the characteristics of national populations. These two forms of knowledge production were oriented to quite different ends: while CBA sought to enable decision-making, even if it meant assigning values to highly intangible costs and benefits, sampling techniques sought accuracy and ‘truth to nature’. The things that counted as important in these settings were quite different – the

production of these kinds of knowledge was oriented to different ends, and strived for different kinds of good quality.¹⁰

The phrase ‘orientations to knowledge production’ is a bit cumbersome, so instead we will use ‘axes of quality’.¹¹ This is the same idea – where there are different purposes for knowledge, there will be different axes of quality. Think, for example, of the prospect of a standardised framework for impact measurement used by multiple different organisations. In this situation, a good quality IMM system would have a comprehensive, overarching set of outcomes and indicators that organisations report against, while detailed qualitative feedback would not be valued or useful. In contrast, an organisation seeking in-depth understanding of the difference it has made to a small number of direct beneficiaries would work with an axis of quality that rates detailed feedback highly, while simple quantitative measures would by themselves be considered low quality. A social impact bond that needs robust outcomes measures would have yet another axis of quality for knowledge about impact, perhaps valuing certainty of attribution.

We can use this idea of ‘axes of quality’ to think more clearly about what we should be aiming for in developing IMM good practice, and to see the inadequacies in our current way of approaching this topic. For an organisation to be able to build a high quality IMM system that is fit for purpose, that organisation needs a clear sense of what counts as good quality, and what they should seek to avoid – they need a clear sense of the relevant axis of quality. Currently, I think we are labouring under the impression that it is possible to define a single account of ‘what good looks like’ for IMM. I think we are assuming that there is a single axis of quality for all IMM activity.

The reality, I am arguing, is that we should set out to

define multiple possible axes of quality within IMM practice, and that they depend on the purpose for which IMM is being undertaken – something that depends on all kinds of factors, such as the audience for the information, whether you are an investor or investee, whether you work directly with end users, and so on.

The notion of ‘standards of evidence’ needs a mention at this point. Pioneered by Nesta in 2012, standards of evidence map out the stages between a basic starting point for generating evidence of impact (being able to articulate what impact will be created by an intervention – “level 1”) through to the most robust, defensible forms of evidence (evaluations that demonstrate replicability of impact in other settings – “level 5”).¹² This is, in fact, an axis of quality. The problem is that it is presented as generally applicable. It reinforces the message that robustness and reducing uncertainty is always best practice, regardless of the particularities of any given attempt to better understand impact. What should an organisation do when they can see from the outset that the methods needed to achieve the higher standards of evidence will never be relevant or appropriate in their case?

Currently – and this is true of the ‘standards of evidence’ approach and more broadly – they are offered an equivocation: this is best practice, the argument goes, but your measurement response also has to be proportional and appropriate to your situation. In other words, organisations are simultaneously told that they should strive for the highest standards and that it is perfectly legitimate if they consider the highest standards to be irrelevant in their case.

We need to find a way to change the conversation.

¹⁰Lengthy discussion of the origins of cost benefit analysis are found in Theodore Porter, *Trust in Numbers: The Pursuit of Objectivity in Public Life* (1995), while the history of accounting is discussed in Ian Hacking, *The Taming of Chance* (1990) and Mary Poovey, *A History of the Modern Fact: Problems of Knowledge in the Sciences of Wealth and Society* (1998).

¹¹This term is not borrowed from anywhere else, and any overlap with the work of other scholars is coincidental.

¹²Nesta, Ruth Puttick and Joe Ludlow, *Standards of Evidence for Impact Investing* (2012). p8

Internal differentiation

If IMM is going to develop as a rigorous and systematic set of practices, we need to find a way to formalise the differences between these axes of quality. This is a challenge faced by any attempt to systematically or scientifically produce knowledge – there will always be disagreements over what counts as good practice, so these disagreements have to be clearly delineated, the foundational differences understood. If we take any discipline that has its own rules for producing knowledge there will be a detailed, lengthy debate over what counts as good practice. For example, Bayesian and frequentist approaches to statistics entail different statistical methods, and these differences continue to be discussed decades later. This is what I mean by internal differentiation – the drive to understand the roots of these differences, and justifications for choosing one option or another. Such debate over the best way to produce knowledge makes it possible for the community to hold each other to account over the results they present.

As things stand in IMM there is an *almost total lack of internal differentiation*. There are few, if any, attempts to consider where different forms of IMM are built on the same axis of quality, oriented to the same ends, and where they are actually trying to achieve different things. There are few, if any, attempts to consider the basic incompatibilities of different approaches to IMM. This is a matter of epistemology. This kind of debate is normally the province of academics, but so far the emphasis has been on IMM as a tool for practitioners, while academic concerns are seen as removed from the everyday realities of running an enterprise or managing capital. While these pragmatic concerns are important, these issues will not go away. If IMM is ever going to develop conventions that allow for meaningful results, and mutual accountability, it will have to confront them.

Internal differentiation, then, is partly about recognising different axes of quality, as described above. It is also about challenging the assumption that understanding impact is simply a matter of collecting ‘im-

part data’ that is ‘analysed’ in a generic manner. It means recognising that before we even reach the stage of collecting data, a huge amount of legwork is needed in specifying a framework – built on a clearly specified axis of quality – that directs us to the most relevant data to collect. It is this framework, furthermore, that allows for meaningful analysis. It has to become far more commonly understood that impact metrics do not speak for themselves – they are given meaning by the frameworks we put them in. These frameworks are built to suit different purposes, with corresponding axes of quality. Once these basic features of knowledge production are recognised it will become possible to see where existing IMM tools and approaches are compatible with the particular needs of an organisation, and where they are incompatible.

It is important to note that there are plenty of IMM approaches that are very clear and explicit in setting out when they should be used, and what it means to implement them well. SROI is a good example of this, as its seven principles of social value provide guidance over what is important in conducting an SROI, and the assurance process provides explicit criteria for success. Similarly, Impetus PEF’s Driving Impact and Acumen’s Lean Data both put forward visions that are grounded in a philosophy of what IMM should be for, and directly tackle questions over what kind of information should be gathered at what stage, and why. In other words, they take time to mark out an axis of quality. These solutions – and there are many more – are valuable contributions to the collective attempt to improve IMM practice.

But even with these clear contributions to the debate, the overall tone is still focused on consensus, on finding shared common ground and diminishing difference and incompatibility. It is rare to find any kind of detailed discussion about the differences between IMM approaches, or the limitations of any particular method.

Why has the need for differentiation not become clear already? Why do we so rarely encounter points of direct conflict over IMM approaches? I can suggest one

reason: each attempt to do IMM well also comes up against the practicalities of producing knowledge about impact while running an organisation, and so the limits of the current logic are never tested. If any given attempt to implement IMM falls a long way short of the ideal 'best practice', then incompatibilities in different approaches are never brought to the surface.

If we can start bringing more attention to *multiple* possible axes of quality, then we can collectively start to create internal differentiation. It will then be a matter of ongoing debate and discussion which axes of quality should be adopted and used as a reference point, and which are unhelpful, *unfit* for purpose.

Changing scenes

Based on this logic, we will need to pay more attention to the *transfer* of impact data between settings that require different axes of quality. This is important because the transfer of impact data between settings increases the risk of miscommunication and misleading statements about impact.

We have seen this already with the NN Investment Partners example: they developed a framework for internal use, then also used it as the basis for making claims in a published article. Impact data has also transferred settings in the NPC/KLF example, as NPC used outcomes data they had found on portfolio companies' websites as contributing to the impact of the KLF Foundation. Whenever the recipients of funding or investment are asked to pass their impact data on to their donor or investor, the same thing is happening: impact data is being transferred from one setting to another.

We remarked above that this is one of the reasons why

transparency over uncertainty is needed. If certainty is not a priority in one setting, then we need to make sure that, on being transferred to another setting, the impact data is used in a way that is not misleading about the degree of certainty in the results.

We can add to this that the assumption of transferability is undermined where there are different axes of quality. If, for example, a social enterprise prioritises detailed, structured qualitative feedback to understand its impact then the knowledge produced might be well suited to their purposes, but be entirely incompatible with their investor's impact measurement framework. Of course, it is entirely possible for an investor and investee to work together to agree a common framework, using a common axis of quality. But I don't think the demands – or the compromises – of agreeing such a framework are sufficiently recognised. The point is that transferability *cannot be assumed* if there is no explicit effort to make such an agreement.

Think back to the earlier example taken from the OECD's report: they recommend "data should be captured and aggregated in an internationally comparable manner". We can now see that this vision can only be brought about if some sort of internationally agreed upon framework is also brought into play – a framework, crucially, that is detailed enough to specify an axis of quality. Current attempts, which leave the 'analysis' element of understanding impact unspecified, do not go far enough. Of course, if a detailed framework is specified 'from above' and then filtered down to a wide variety of situations, there will be significant compromises on how relevant the framework will be in any given setting. By recognising these compromises from the outset, we can make better informed decisions about what to aim for.

The bigger picture

Before we go any further, there is a hugely important aspect to the IMM debate that needs bringing into play: the global narrative around impact investing. It has become taken for granted that IMM is necessary and

important. Why? Because impact investing relies on IMM for its continued credibility. Without IMM, impact investing is just investing as usual. This is a strong claim, and I unpack it in more detail elsewhere. The

point I want to make here is that the debate about the need for IMM, and the purposes for which it is being developed, is heavily influenced by the perceived need to protect against impact washing, and to ensure ‘transparency and accountability’.

All of these concerns take shape in the context of the attempt to build a market for impact investing. The market provides the framework for impact investing activity. Energy is being poured into this market building agenda: how do we maximise the flow of capital into this market? How do we bring supply and demand together? How do we reduce the barriers present in the market? How do we improve the investment readiness of impact-driven organisations? And so on. All of this has a significant effect on how we think about IMM. After all, IMM systems have the potential to help accelerate the growth of the market, or to hinder it.

If IMM, as a rule, were an intensive, lengthy, expensive procedure, it would act as a barrier to entry to this market. If, in contrast, IMM practice was generally light-touch, producing information that is easily understood and transferable across settings, it would be highly compatible with the operation of a market, and add to the information flows that already take place. I will not attempt to analyse the reality of this situation – it is far too complex to tackle here – but rather point to the probable effect of the market building agenda on the way we collectively think about what IMM should look like.

If we are aware of this effect, then we are better placed to weigh up the compromises involved.

After all, if we were building an IMM system to help the growth of the impact investing market, we would be using an axis of quality that is quite different to the axis of quality underpinning detailed, local knowledge of impact. Before we continue the monumental effort of building standardised frameworks for comparable impact data, as the OECD suggests, we should recognise that these systems will meet some purposes and not others, and decide collectively what is most important, and which compromises should be made. This only becomes more pressing when we connect it

up with the question of cause, effect and uncertainty. The question is not just what compromises have to be made on compatibility of different axes of quality, but also on the degree of uncertainty that is appropriate or desirable. Would we want a standardised, internationally comparable IMM framework if it produced impact knowledge with huge amounts of uncertainty?

We can return to look at the NPC/KLF report in light of this bigger picture. Indeed, the introduction to the report states

If we are to ensure that this growing pool of assets has measurable and relevant impact—and as a result is able to continue to attract further flows of capital—then it is vital, now more than ever, that we do our very best to assess the social and environmental outcomes of those investments and get a handle on the likely impact of that capital.

In pursuit of deep impact, p4

To my mind, this scene-setting tells us a lot about the motivations behind this report and the conclusions it draws. Despite the numerous caveats and issues that are set out in the body of the text, NPC still claim in the opening paragraphs that “despite challenges for investors and investees, it is possible to measure the outcomes of a wide spectrum of investments across different asset classes and impact categories”, and that “impact (of varying degrees) can be achieved while gaining a financial return (of varying levels)” (p4). These findings smooth over the difficulties they have encountered, which we can now see stem from enormous amounts of uncertainty in the data they had available, and from assuming transferability of data between settings without a shared axis of quality.

I am suggesting a change of perspective. This argument runs somewhat against the grain of the current debate. If I can convince the reader, it means a reassessment of what is possible, of the long-term aims for IMM in the context of impact investing, and of the demands of doing IMM well. It is my view that this reassessment is

necessary if IMM is going to generate meaningful insight into impact.

So where do we go from here?

What should be done?

This paper does not offer a simple solution. There is no framework or typology here. Equally, it does not contain sets of criteria for what should be seen as good or bad IMM. That is for others to decide, in relation to their own needs and constraints. Indeed, I hope to convey the message that such attempts to generalise are unhelpful, and should be put aside in favour of recognising that different circumstances require different approaches.

But what will it mean to be ‘more transparent about uncertainty’? What will it mean to be ‘clearer about the axis of quality’ in a particular setting? What will it mean to be realistic about the ‘transferability of impact data’? I would offer a straightforward answer to these questions if I could, with accompanying diagrams and categorisations, but the reality is these matters need to be tackled in relation to each particular setting. In turn, tackling them requires judgement and a foundational level of expertise. To elevate the debate around IMM, we need to ensure that participants have the ability to judge the degree of uncertainty, to balance different priorities, to understand the unavoidable compromises of choosing one IMM system over another, and to compare these decisions with those of others, and respond to challenges.¹³

By calling for expertise I do not mean to put IMM out of reach, to make it the domain of an elite IMM profession. Not at all. In fact, I would argue that the current degree of confusion and ambiguity mean this is where we are already at, despite all attempts to make IMM accessible.

The kind of expertise I am talking about is not impossible to acquire, and is already present in our commu-

nity. It is the bread and butter of neighbouring industries, particularly the monitoring and evaluation profession, which has been working with these ideas and concepts for decades. It is familiar to any social science student. At root, it is nothing more than standard research methods training. In this sense, the change I am talking about is a relatively small shift towards bringing in those who are trained in the challenges of producing knowledge about the social world. Asking somebody to produce knowledge about impact without this training is like asking somebody with no financial training to conduct due diligence on a company. Just like finance has a set of conventions and concepts that have to be learned before judgement can be applied, so does IMM. The basics of IMM are understanding cause, effect and uncertainty, and being able to differentiate axes of quality.

In another sense, this is a monumental shift. There is a lot of work to be done. I have talked about axes of quality as if we could lay them out side by side, each with its own set of clearly defined criteria. This would be too literal an interpretation. The idea of an ‘axis of quality’ is better understood as a metaphor that helps to communicate that what counts as good quality varies from one situation to the next. Even within a single, well-defined setting, there are still multiple factors to take into account, and getting an IMM system to work well in capturing the messy, chaotic reality of people’s lives is never easy. There will always be gaps in the data and things we do not know. Just like someone’s ability to do financial due diligence improves with experience, so does the ability to do IMM.

Spend some time reading about the state of monitor-

¹³In this sense, my argument is highly compatible with the work of Kate Ruff and Sara Olsen on the potential benefits of bringing the skills of evaluation analysts into impact measurement. Kate Ruff and Sara Olsen, ‘The Need for Analysts in Social Impact Measurement: How Evaluators Can Help’, *American Journal of Evaluation*, 39/3 (2018)

ing and evaluation in international development and it will become clear that there are no straightforward answers to the question of what counts as good quality. I imagine some readers of this paper will sigh in frustration at the suggestion that we deliberately open up such topics of debate in the context of impact investing. To this I respond: these debates are a feature of the attempt to create meaningful insight into the nature of social change. The only way to avoid the question of cause, effect and uncertainty is to ignore it, which does not mean it goes away – instead, it undermines the endeavour to create meaningful insight. So far, the im-

plications of avoiding uncertainty have been obscured by the general degree of confusion, and insistence on building consensus instead of recognising difference.

Collectively, we need to get better at putting boundaries on our attempts to produce knowledge about impact, and at adjusting out expectations over what is possible. We can then move past the current status quo, where we are constantly striving for an ideal that we know we cannot reach, and that we are unsure is even what we want.

J Dagers
May 2019