

# Evaluating policy organisations

## Summary

This document outlines some of the crucial considerations to take into account when evaluating charities that advocate for political change. There is little published research on the best way to approach this problem, so our conclusions here are tentative and we expect them to change as we refine our approach over the next few years.

In this document, we discuss:

1. How to evaluate the counterfactual benefits of a policy campaign. Key factors include:
  - The chance that a similar campaign would have been run later.
  - The chance that a law will be repealed.
2. How to evaluate the counterfactual difference an organisation made. Key factors include:
  - The importance of testimony.
  - The merits of gathering primary and secondary evidence.
  - Whether organisations should get a portion of ‘the credit’ for a policy success.
  - The importance of coordination among funders and organisations.
3. The merits of building quantitative models to assess the impact of policy campaigns.
4. The chance of success with policy advocacy. Key issues include:
  - Policy impact is plausibly distributed according to a power law: most policy campaigns fail, but there are some big wins, which make a portfolio of bets worthwhile.
  - The need for patience when funding campaigns.
5. How to evaluate future impact. Key factors include:
  - What we can learn from the track record of a policy organisation.
  - The importance of project selection for future impact.
  - How well the organisation’s activities fit in with a plausible theory of change.

## 1. Introduction

Government policies can have a large effect on the welfare of huge numbers of people. Policy advocacy therefore constitutes an exceptional opportunity for leverage: by shifting government spending or changing the law, non-profits can improve the lives of millions of people. Even though the probability of success of any policy campaign is small, the rewards of success are so large that the risk is justified. How, then, should we go about assessing policy advocacy campaigns?

It is first useful to clarify some important background concepts. Donors trying to do good should aim to have *counterfactual impact*, which is the difference between what happens as a result of your donation and what would have happened otherwise. From this perspective, it is not sufficient to establish that Charity A led the campaign that changed government policy; we also need to ask what would have happened if they had not led the campaign. Perhaps if Charity A hadn't existed, one of a dozen similar alternatives would have stepped in almost immediately to lead an identical campaign. If so, the impact of Charity A's campaign was at least much lower than it first appears.

Indeed, policy interventions often seek to redirect government resources from less to more effective programmes. This means that, most of the time, the beneficiaries would have enjoyed some, although less, benefit, even without the organisation's intervention. When considering the effect of the policy intervention, therefore, we should not simply calculate the benefit brought about by the new policy. Rather, we need to consider the benefits brought about counterfactually—that is, the difference between the benefits that did in fact materialise and those that would have occurred had the organisation not intervened.

Another key concept for assessing policy advocacy is *expected value*. It will always be uncertain whether a political campaign will succeed, and when acting under uncertainty, we should aim to maximise the expected benefit of our donation, which is given by the following equation:

$$\text{Expected benefits} = (\text{probability of success}) * (\text{benefits of success})$$

('Success' here refers to making a counterfactual difference.) However, figuring out the value of these two factors can be very difficult. Even retrospectively assessing the expected benefits of a past campaign tends to involve large uncertainty, and assessing the expected benefits of a future campaign is usually even more uncertain.

In this short brief, we will discuss some of the key factors relevant to the evaluation of policy campaigns.

## 2. Evaluating the benefits of policy change

Understanding the benefits of a given policy change is a crucial component of evaluating the expected benefits of a policy campaign. In some cases, calculating the benefits of a policy is relatively straightforward because the policy has been assessed by reputable organisations in publicly available reports.<sup>1</sup> In other cases, we will have to calculate the benefits using other available data.

For people concerned about impact, it is most important to consider the benefits of an action compared to what would have happened otherwise. The total counterfactual benefit of a policy campaign is given by:

**Total counterfactual benefit** = (benefits brought about by campaign) – (benefits that would have occurred without the campaign)

In light of this, two key points to consider when assessing the benefits of a policy change are that:

- A successful policy campaign usually brings the implementation of a policy forward in time, i.e. it doesn't make the difference between the law happening and never happening.
- There is a chance that the law will be repealed in the future.

To illustrate the importance of these two points, we use an example adapted from our 2018 climate change report, involving the Clean Air Task Force (CATF), a US-based non-profit that advocates for low carbon energy.<sup>2</sup> In the 2000s, the CATF ran a successful campaign for the Obama administration to regulate methane emissions from natural gasworks. How should we calculate the benefits of this policy change?

### Bringing policies forward in time

One might think the total benefit of CATF's campaign is given by the following equation:

**Total benefits** = (average annual benefits of the regulation) \* (time that the regulation is in place)

However, this calculation would be incorrect because we need to consider the counterfactual; we need to ask what would have happened if CATF had not run its campaign. In all likelihood, if CATF hadn't run its campaign, another NGO would, later, have stepped in and advocated for a similar regulation at a later date. Thus, the counterfactual effect of CATF's campaign is probably to bring forward this beneficial policy by some amount of time.

---

<sup>1</sup> This is true of many of the EPA regulations advocated for by the Clean Air Task Force.

<sup>2</sup> Available at [www.founderspledge.com/research](http://www.founderspledge.com/research).

### Potential repeal of policies

The other crucial aspect of policy assessment is the probability that the policy will be repealed. Suppose that the Trump administration is hostile to environmental regulation, so we estimate there to be an 80% chance the regulation will be repealed at the end of 2019. Thus, we should discount the benefits of the regulation by the probability that it will be repealed in a given year. In this case, the benefits of the regulation are given by:

$$\text{Benefits of the regulation} = (\text{benefits of the regulation in 2015 to 2018}) + (\text{benefits of the regulation in 2019}) \cdot (1 - 0.8)$$

Thus, we always need to consider the chance that a given policy will be reversed in the future. This will usually depend on the level of opposition to the policy among the public and political leaders. The policy might also have a specific end point.

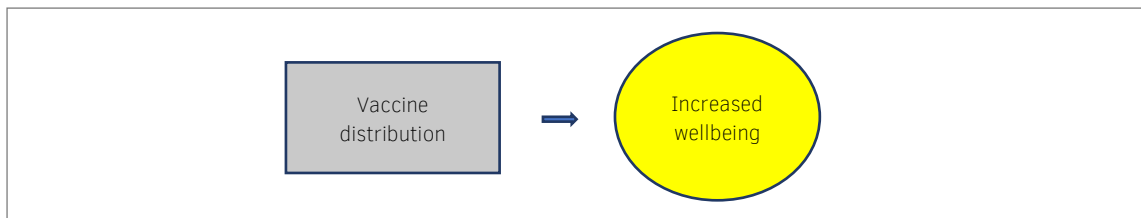
### 3. Evaluating the counterfactual role of an organisation

Our second task when evaluating the expected value of a campaign is to work out the counterfactual role played by the organisation. This raises a number of practical and philosophical issues.

#### Who made a difference?

One of the features that distinguishes policy interventions from other types of philanthropic projects is their highly indirect nature. Many interventions supported through philanthropy consist of interventions that directly bring about impact. For instance, by distributing vaccines, bed nets, or cash, an organisation can fairly directly bring about improvements in the wellbeing of the recipients, as shown in Figure 1:

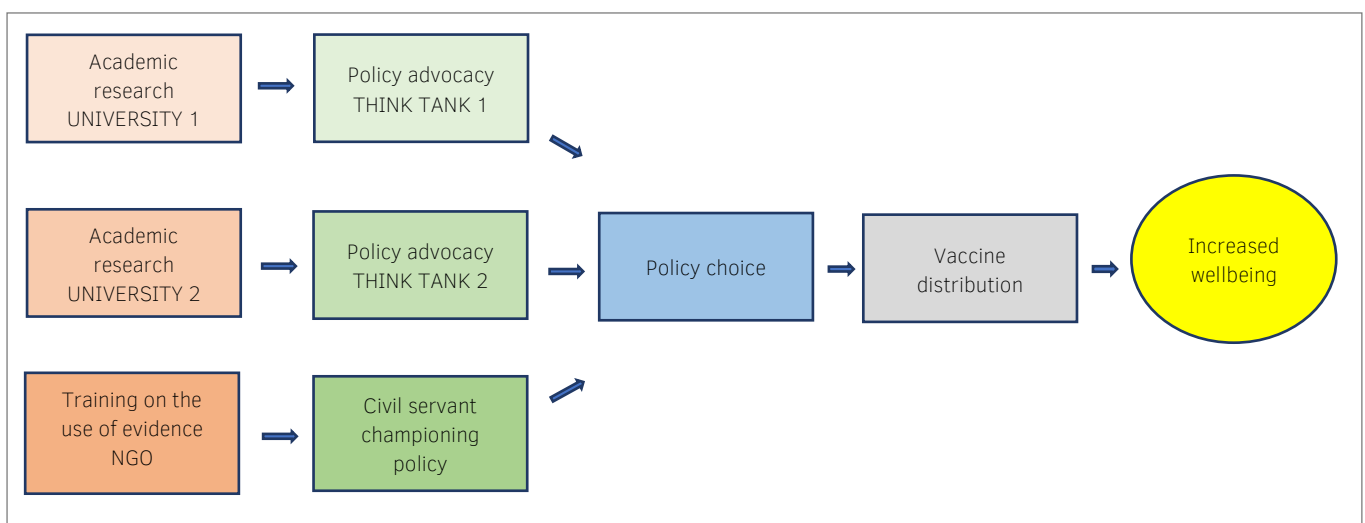
Figure 1. A direct intervention



For charities doing direct work, it is often relatively straightforward to work out who brought about a particular change. For example, we know with high confidence that GiveDirectly is responsible for the delivery of a certain amount of unconditional cash transfers in Sub-Saharan Africa.

Policy interventions, in contrast, are much more indirect and typically involve many more actors: there are more steps on the path from the charity to the desired outcome, as shown in Figure 2:

Figure 2. An indirect intervention



These characteristics of policy work make it hard to figure out the counterfactual effect of each actor. If there are 50 NGOs advocating for a policy change, how do we know which one

made a genuine difference? How replaceable was each organisation: if one of them had not joined the campaign, would another group have stepped in? How overdetermined was the event: were there numerous campaigns each advocating for the same change? Was the campaign even relevant: would the government have implemented the policy anyway?

It is very hard to answer these questions. Consequently, when assessing policy, we must rely on much weaker evidence than we would when assessing charities doing direct work. At best, we will be able to make probabilistic statements like “without Clean Air Task Force, there is a 20% chance that the law would have been delayed by two years”.

### Gathering evidence

This suggests that it is crucial to gather the following evidence when assessing counterfactual impact:

- **Crowdedness of the field** – If there are numerous other organisations with similar aims and capabilities, this suggests that the work done by the organisation is more replaceable, i.e. another organisation would have readily stepped in to do the work.
- **Role of each actor** – It is important to figure out the role of each actor involved in the campaign in order to understand who played a necessary role in achieving a particular outcome.
- **Consistency of timelines** – Does the timeline of impact provided by the charity fit with different possible timelines of when decisions were made in government?
- **Catalytic nature of the charity’s work** – Conceiving and leading a campaign is often less replaceable than joining a campaign. Therefore, if an organisation played a catalytic role, the charity probably had greater counterfactual impact.
- **Nature of government stance** – Is there evidence that the government would have made the change anyway even in the absence of an advocacy campaign?

There are two different promising broad approaches to evaluating advocacy:

1. Evaluating secondary evidence of the charity’s role in a particular policy change.
2. Evaluating the primary work of a charity in a campaign.

Certain aspects of these two approaches may often be combined, but it is useful to compare them in isolation.

#### *Evaluating secondary evidence*

On the first approach, there are two steps involved in reconstructing the counterfactual impact of the organisation. Firstly, we gather narrative information from the charity outlining their role in a policy success, an account of how the change was made, and indications of their replaceability in the campaign.

Secondly, we gather evidence from independent referees and written sources to confirm or disconfirm the charity’s own account of their impact. There is a hierarchy of testimony evidence, ranked from the most to least desirable:

1. Well-informed people with incentives to downplay the role played by the organisation:
  - a. A rival organisation who were deeply involved in a parallel campaign with similar aims, who would wish to claim the credit for themselves.<sup>3</sup>
  - b. Politicians involved in the decision but opposed to the organisation.
  - c. Companies who have been the subject of reputation-damaging advocacy campaigns.
2. Well-informed people with no incentive to mislead about the role played by the organisation:
  - a. Government bureaucrats or politicians who were directly involved in the policy change.
3. Well-informed people who have incentives to mislead about the role played by the organisation:
  - a. Grantmakers who supported the organisation.
  - b. An organisation's campaign partners.
4. People with less information on the role played by the organisation:
  - a. Short-form media sources that report on the policy change.
  - b. Government bureaucrats or politicians who were not directly involved in the policy change.
  - c. Advocates for the policy change that did not play a role in bringing it about.

On this approach, when we are assessing a past policy advocacy campaign, we usually ask charities for two to three referees to vouch for the role played by the charity.

Other useful sources include long-form political histories in books, newspapers, or blogs. For example, Seymour and Busch's *Why Forests, Why Now?* contains a lot of detail on the role played by one of our climate change charities, the Coalition for Rainforest Nations, in efforts to prevent tropical deforestation.<sup>4</sup>

Testimony evidence can also be useful for assessing the future work of an organisation. Impartial experts can provide evidence on whether a future campaign is likely to succeed. For example, an impartial government bureaucrat could provide information on the amount of influence that a particular NGO has in government, or on whether leading figures in government are receptive to changing a policy.

---

<sup>3</sup> Note that not all parallel campaigns will necessarily be competitive or antagonistic.

<sup>4</sup> Frances Seymour and Jonah Busch, 'Why Forests? Why Now? A Preview of the Science, Economics, and Politics of Tropical Forests and Climate Change', Center For Global Development, November 2014, <https://www.cgdev.org/publication/ft/why-forests-why-now-preview-science-economics-politics-tropical-forests-climate-change>. See our climate change report at [www.founderspledge.com/research](http://www.founderspledge.com/research)

### *Evaluating primary work*

Another possible approach to evaluating the role played by an organisation is outlined in [The Elusive Craft of Evaluating Advocacy](#) by Teles and Schmitt:

“The best way to evaluate an organization whose influence is extremely diffuse is for grant officers to be close to the political action and thus able to make informed judgment calls on how it conducts its core activities. This was the practice of many conservative foundations, whose staff devoted much of their time to simply reading the primary work of their grantees, rather than asking them to generate problematic metrics and lengthy reports designed solely for purposes of evaluation. Empowered by their boards or donors to trust their own judgment of good, appropriate work, this foundation strategy has been vindicated many times over in the real world of politics and the marketplace of ideas.”<sup>5</sup>

This approach requires grantmakers to be close to the political action and to have a deep understanding of the role played by each organisation on the ground.

Each of these two approaches has advantages. The first approach can be burdensome on charities, making them less likely to join the evaluation process and diverting resources from their core mission. Moreover, asking charities to produce reports on themselves incentivises gaming of artificial metrics. The reports are also likely to be an inferior source of information on the activities of the organisation than knowledge of their primary activities.

The main downside of the second approach is that it requires full-time specialised grantmakers to be embedded in a particular area over the long term. For charity evaluators without the capacity to do this, the second approach may not be viable. The first approach is more suited to a charity evaluator engaging with a field for less than a year.

### **Leveraging and funging**

Because policy advocacy usually involves multiple actors, it is important to bear in mind interdependencies between these actors when assessing impact. Policy campaigns almost always depend on the efforts of a wide range of actors. Sometimes, organisations will only join if numerous other organisations are involved. Other times, an organisation merely displaces another organisation from a campaign.

More technically, we need to be aware of *leveraging* and *funging* effects:

**Leveraging** – The resources spent by one group on a given problem cause another group to spend more resources on that problem.

---

<sup>5</sup> Steven Teles and Mark Schmitt, ‘The Elusive Craft of Evaluating Advocacy’, *Stanford Social Innovation Review*, 2011, <http://familyplanning.care2share.wikispaces.net/file/view/Elusive%20Craft%20of%20Evaluating%20Advocacy.pdf/348250912/Elusive%20Craft%20of%20Evaluating%20Advocacy.pdf>.



**Funging** – The resources spent by one group on a given problem cause another group to spend fewer resources on that problem.

Here is an example of leveraging: suppose that because the Against Malaria Foundation (AMF) has committed to spend \$1 million on malaria bed net distribution, the Gates Foundation contributes \$500,000 to the distribution. In this case, AMF leverages the resources of the Gates Foundation into the problem of malaria. AMF's counterfactual impact is then given by:

AMF's own \$1 million on bed nets *plus* the Gates Foundation's \$500,000 on bed nets *minus* the benefits of what the Gates Foundation would otherwise have spent their \$500,000 on.

Here is an example of funging: suppose that AMF has committed \$1 million to a bed net distribution scheme, but if they had not done this, then USAID would have committed \$500,000 to the distribution. In this case, AMF funges with USAID. AMF's impact is then given by:

AMF's own \$1 million on bed nets *minus* the \$500,000 that USAID would have put in *plus* the benefits of what USAID in fact spent their \$500,000 on.

Quantifying leveraging and funging effects is difficult because it requires modelling a highly uncertain counterfactual. It is nonetheless important to take account of these effects.<sup>6</sup>

### The paradox of double counting 'credit'?

On the account we have outlined in this section, the counterfactual impact of an organisation is given by the difference between the following worlds:

**Actual World** – The world in which the organisation acts.

**Counterfactual World** – The world that would have been brought about if the organisation had not acted.

If there is an additional \$1 billion in social benefit in the Actual World compared to the Counterfactual World, then the counterfactual impact of the organisation is \$1 billion. But this gives rise to an apparent paradox. Suppose that there were two organisations set up to campaign for a policy change that produced a \$1 billion benefit, Braintank, which specialises in research, and Actionlab, which specialises in lobbying politicians. Each organisation is necessary for the campaign's success: if either one of these two organisations had not been formed, the campaign would have failed. And each organisation acts independently: if either one had not tried to campaign, the other organisation would have campaigned regardless.

---

<sup>6</sup> Our research partner GiveWell now accounts for leveraging and funging effects in its [cost-effectiveness analyses](#).

So, by the method outlined above, Braintank's counterfactual impact is \$1 billion and Actionlab's counterfactual impact is \$1 billion. Doesn't this imply that the combined counterfactual impact of Braintank and Actionlab is \$2 billion? If so, since this is greater than the total benefits of the policy, the method we used to calculate the counterfactual impact of each organisation must be wrong. It seems that we have erred in giving each of Braintank and Actionlab 100% of 'the credit' for the benefits. Perhaps they should each get 50% of the credit, so that each is responsible for \$500 million.

However, in our view, there is no paradox here: the method for calculating counterfactual impact is correct. All this example shows is that we cannot always aggregate the counterfactual impact of different agents, not that the conception of counterfactual impact is wrong. There are two main arguments for this. Firstly, the definition of counterfactual impact seems to express what the basic concept means, and we are not sure how to define it in a way consistent with the alternative credit-based view. If we can, as argued below, retain this concept without sacrificing anything of intuitive importance, then there is reason to do that.

Secondly, more closely examining the meaning of statements about counterfactual impact shows why we cannot unproblematically aggregate counterfactual impact. When we are assessing the *combined* counterfactual impact of Braintank and Actionlab, we are comparing two worlds:

**Actual World** – Braintank and Actionlab are formed and complete their successful campaign.

**Counterfactual world<sub>A&B</sub>** – The world in which neither Braintank nor Actionlab are formed.

As would be predicted by our definition of counterfactual impact, this implies that the combined counterfactual impact of Braintank and Actionlab is \$1 billion, which is intuitively correct. This counterfactual comparison is different to the one that we get when we aggregate the counterfactual impact of Braintank and Actionlab, each taken individually. Braintank's counterfactual impact is given by the difference between:

**Actual World** – Braintank and Actionlab complete their successful campaign.

**Counterfactual world<sub>B</sub>** – The world in which Braintank does not exist and Actionlab acts as it would have had Braintank not existed, i.e. it unsuccessfully campaigns for the policy.

Similarly, Actionlab's counterfactual impact is:

**Actual World** – Braintank and Actionlab complete their successful campaign.

**Counterfactual world<sub>A</sub>** – The world in which Actionlab does not exist and Braintank acts as it would have had Actionlab not existed, i.e. it unsuccessfully campaigns for the policy.

If we try to aggregate these two counterfactual comparisons, then we are in effect making this comparison:

**Actual World** – Braintank and Actionlab complete their successful campaign.

**Counterfactual world<sub>A+B</sub>** – The world in which Braintank does not exist and Actionlab acts as it would have had Braintank not existed, i.e. it unsuccessfully campaigns for the policy; and Actionlab does not exist and Braintank acts as it would have had Actionlab not existed, i.e. it unsuccessfully campaigns for the policy.

This is an incoherent comparison: in the posited counterfactual world, Braintank both exists and does not exist and Actionlab both exists and does not exist. This illustrates that the aggregation principle is wrong: we cannot always reliably aggregate the impact of agents, taken individually. Since the aggregation principle is wrong, there is no reason to get rid of the conception of counterfactual impact.

This also shows that there is no reason to assign policy organisations a percentage of ‘credit’ for a policy change that must add up to 100%. It is not clear what ‘credit’ means in terms of counterfactual impact, and given that aggregated counterfactual impact does not have to add up to 100%, the motivation for the notion of credit appears lacking.

Indeed, if we understand counterfactual impact as credit, we will be led to make mistaken decisions in some cases. Suppose that David is one of a billion voters in an election to decide on a law that will produce a \$5 million benefit to the economy. The actions of each of the billion voters are independent—they will vote regardless of what any of the others do. David knows that his vote will be decisive: if he votes, the law will be passed and if he doesn’t vote, it won’t. Suppose that David’s cost of voting is \$0.01. If David believes the conception of counterfactual impact that we have outlined above, then he will reason as follows: “I have to pay \$0.01 to produce a \$5 million benefit, a 1:500 million cost-benefit ratio, so clearly I should vote.”

However, if he uses the notion of credit, then he will say “the credit for the possible \$5 million benefit must be divided between one billion people, so my counterfactual impact is \$5 million/1 billion, or \$0.005. So, I would have to pay \$0.01 to produce a benefit of \$0.005, a 1:0.5 cost-benefit ratio, so I will not vote.” Reasoning in this way, David abstains, everyone else votes, and the law is not passed, so the economy loses out on \$5 million. David has made a mistake here, which is driven by his acceptance of the notion of ‘credit’.

### **Counterfactual impact and coordination**

The foregoing argument shows that it is important to be careful when aggregating the counterfactual impact of organisations working on the same campaign, and it is important to consider how organisations working on the same problem can coordinate. Consider a decision facing a prospective funder of the campaign involving Braintank and Actionlab. The counterfactual impact of each, taken individually, is \$1 billion, but it would be a mistake for the

donor to add this counterfactual impact up and conclude that the total impact of funding both is \$2 billion. If the campaign costs \$1.5 billion, then the campaign would not be worth it, but it would be easy to miss this fact by aggregating the impact of the two groups.

Related to this, if the campaign costs \$1.5 billion in total, it would be optimal for funders of Braintank and Actionlab to coordinate and not run the campaign. Similarly, for the David voting example, although it is worthwhile for David to vote given how everyone else acts, the total cost of all the votes is \$10 million versus a benefit of \$5 million from the policy. In this case, it would be better if everyone could coordinate and abstain from voting. Nevertheless, since coordination is impossible, it is still worthwhile for David to vote.

For donors who are able to fund a full campaign involving multiple organisations, it makes more sense to consider the counterfactual impact of the campaign as a whole when making funding decisions, rather than considering the impact of each organisation involved in the campaign separately. Taking the latter approach unnecessarily introduces complex leveraging and funging questions. In the same way, if you were making a decision about whether to build a factory, you wouldn't evaluate the cost-effectiveness of each part of the factory separately and then try to aggregate the effect, while accounting for leveraging and funging across each part.

Many donors may only be able to fund part of a campaign. These donors should try to coordinate with others as much as possible, but in this case, following the approach of evaluating each organisation independently may be more useful.

## 4. The role of quantification

Quantifying the impact of policy advocacy campaigns is, for the reasons mentioned above, very difficult. In many cases, an evaluation can at best provide a rough indication of the impact of a policy organisation: it will, for example, be very difficult to tell whether a campaign brought a policy change forward by a few months, a few years, or, sometimes, a few decades. The alternative to quantified models is to use intuition to rank organisations and to use verbal descriptions rather than numbers. Quantification has a number of advantages over this alternative approach:

### 1. Quantification conveys information; verbal descriptions do not

Even if you give very wide confidence intervals, such as that the impact of an organisation is between \$1 million and \$100 billion in social benefits, this conveys information about impact. If the widely accepted evidence suggests that the impact of the organisation is within this confidence interval, then by making this quantified estimate, we have made progress versus the world in which we refuse to quantify. Using verbal terms, such as 'huge', to describe an effect also seems inferior because such terms involve significant ambiguity: by being unspecific, verbal models simultaneously afford many interpretations from among which any reader can choose his or her favourite.<sup>7</sup> Evidence suggests that people interpret natural language terms in a huge variety of ways,<sup>8</sup> such that using verbal descriptions can be highly misleading.

### 2. Quantification clarifies the source of disagreement

Related to the above, quantified models make explicit the factors that determine the output of the model. This makes it possible to have productive disagreements about these factors. In contrast, with judgements about intuitions expressed in verbal terms, the source of any disagreement is often not clear.

### 3. Very rough quantification is often possible and can guide decisions

One argument for refusing to quantify or using verbal descriptions is that quantification in many cases seems extremely difficult. However, it is often easier to quantify than one might think at first because larger questions can be broken down into smaller, easier-to-quantify sub-questions. For example, it at first seems difficult to estimate the number of piano tuners in the San Francisco phonebook without much prior information, but in fact with only limited background information, it is possible to

---

<sup>7</sup> Paul E. Smaldino, 'Models Are Stupid, and We Need More of Them', *Computational Social Psychology*, 2017, 317.

<sup>8</sup> M. Granger Morgan, 'Use (and Abuse) of Expert Elicitation in Support of Decision Making for Public Policy', *Proceedings of the National Academy of Sciences* 111, no. 20 (20 May 2014): 7176–84, <https://doi.org/10.1073/pnas.1319946111>.

produce an estimate that is in the correct order of magnitude.<sup>9</sup> Even rough unrealistic, models can therefore help to compare different options and to make decisions.

#### 4. Intuition is flawed

When we are deciding which organisations to donate to, we want to know which organisation will do the most good. The alternative to relying on models to quantify which produces the most good is to use intuitions about which is better. However, there is a wealth of literature showing that intuition is very unreliable.<sup>10</sup> The most important bias for our purposes is *scope neglect*: the tendency for people's preferences to be insensitive to huge differences between different options.<sup>11</sup> For example, there is evidence showing that even if the impact of events differs by two orders of magnitude, people's willingness to prevent those events does not change. Using quantified models is a useful way to protect against scope neglect.

For these reasons, we think it is often worthwhile building quantified models, however rough, to estimate the impact of policy organisations. At their most minimal, these models can be seen as sanity checks of intuitions we have about organisations. Quantification is also often useful for selecting between interventions—the scale and crowdedness of a problem or an intervention are often quantifiable and are a useful guide to prioritisation.

Quantification might not be necessary if one organisation dominates another in terms of features that contribute to effectiveness. For example, suppose that two organisations both work on a high-impact problem and both have a strong team, but one has a much better track record. In this case, qualitative judgements about these considerations are sufficient for deciding which organisation to support. Quantitative models are more useful when each organisation is better than the other in different dimensions: in this case, it will be useful to quantify the size of these differences.

However, quantification involves risks. Firstly, careless quantification may be worse than using intuition; we should not always follow a quantified model over less formal approaches. The argument here is not for following quantified models in all cases, rather it is for the use of careful and judicious quantification.<sup>12</sup> Secondly, quantification may give a false impression of precision and confidence. We might have low confidence in a particular point estimate of the

---

<sup>9</sup> Philip M. Anderson and Cherie Ann Sherman, 'Applying the Fermi Estimation Technique to Business Problems', *The Journal of Applied Business and Economics; Thunder Bay* 10, no. 5 (March 2010): 33–42.

<sup>10</sup> See, for example, Daniel Kahneman, *Thinking, Fast and Slow*, 1st edition (New York: Farrar, Straus and Giroux, 2013).

<sup>11</sup> Daniel Kahneman et al., 'Economic Preferences or Attitude Expressions?: An Analysis of Dollar Responses to Public Issues', *Journal of Risk and Uncertainty* 19, no. 1–3 (1999): 203–35.

<sup>12</sup> Brian Tomasik, 'Quantify with Care', *Essays on Reducing Suffering*, accessed 12 December 2018, <https://reducing-suffering.org/quantify-with-care/>.

impact of a charity, but there is a tendency for outside parties and even for evaluators themselves to start taking such estimates overly literally.

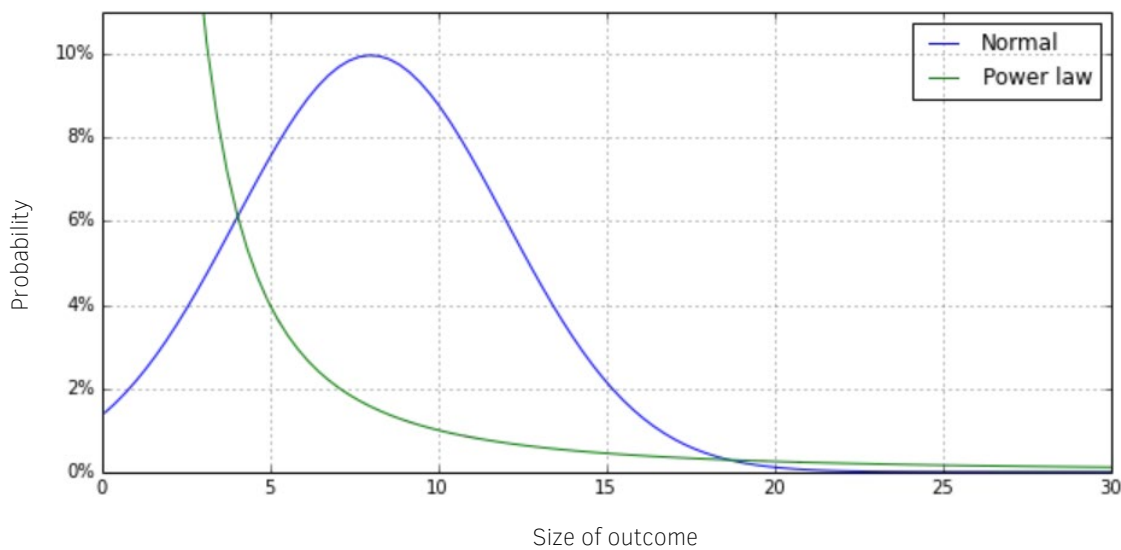
## 5. Big wins, risk and indirectness

We should expect most policy advocacy campaigns to fail. However, this doesn't mean that policy advocacy campaigns are a bad idea because average performance can be pulled up by a handful of 'big wins'.

### Normal versus fat-tailed distributions

The reason big wins are so important for policy advocacy is that the benefits from policy campaigns follow a *power law distribution* with a fat right tail. Other things, such as height and IQ, are *normally distributed*. Figure 3 depicts these two distributions:

Figure 3. A normal distribution and a power law distribution



Source: Jerry Neumann, 'Power Laws in Venture' (2015)

Two things are notable about normal distributions. Firstly, the mean and the median are the same. So, since the height of American women is normally distributed, the total height of American women divided by the number of American women (the mean) is 5 foot and 4 inches, and the typical (median) woman is 5 foot 4. Secondly, in normal distributions, extreme events are rare. 95% of American women are within two standard deviations of the mean height—between 4 foot 10 and 5 foot 10, and it is highly unlikely that we will find a woman who is several standard deviations from the mean. For instance, we will never encounter a woman who is 23 standard deviations from the mean, or 11 foot tall.<sup>13</sup>

Not all events follow a normal distribution. For example, the daily standard deviation of the stock market is around 1%. Despite that, on October 19, 1987, prices fell by 23%—23 standard deviations from the mean—the equivalent of seeing an 11-foot-tall woman. In power law distributions, extreme events are much more likely. Figure 4 zooms in on the right-hand

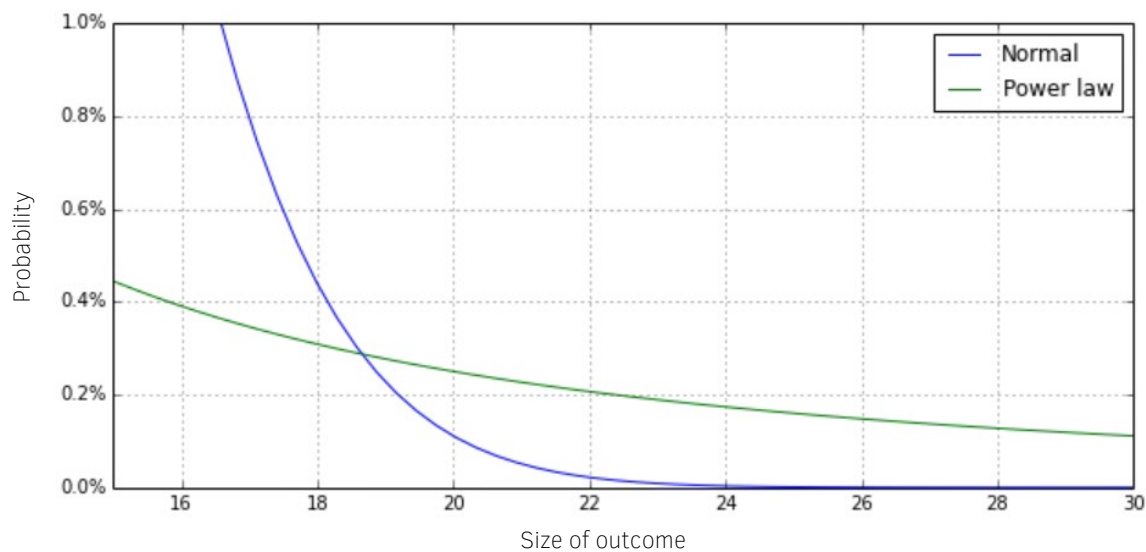
---

<sup>13</sup> William Nordhaus, 'The Economics of Tail Events with an Application to Climate Change', *Review of Environmental Economics and Policy* 5, no. 2 (1 July 2011): 240–57, <https://doi.org/10.1093/reep/rer004>.



tail of normal and power law distributions, showing the frequency of extreme events are in the different distributions:

*Figure 4.* The right-hand tails of a normal distribution and a power law distribution



Source: Jerry Neumann, '[Power Laws in Venture](#)' (2015)

Because extreme events are much more common in power law distributions, a high proportion of the total impact comes from a handful of big wins. For example, most of the returns in a VC portfolio will be driven by a handful of unicorn companies, such as Uber and Klarna. Even though the typical VC investment will fail, the whole portfolio will be driven up by these big wins. Various other things, including the magnitude of earthquakes, the size of cities, the death toll in wars, and so on, follow a power law distribution.<sup>14</sup>

We find it plausible that the impact of philanthropy also follows a power law distribution. By this we mean that, before an organisation runs a campaign, we will not know where the campaign will lie on the power law curve: it could be one of the many campaigns that fails (to the left of the curve) or it could be one of the few big wins (to the right of the curve).<sup>15</sup> There is some support for this from the history of philanthropy:<sup>16</sup>

---

<sup>14</sup> Jerry Neumann, 'Power Laws in Venture', *Reaction Wheel* (blog), 25 June 2015, <http://reactionwheel.net/2015/06/power-laws-in-venture.html>.

<sup>15</sup> There is some evidence that the impact of direct interventions in global health also follow a power law in the sense that across the interventions that have been tried, many have little impact, and some are highly impactful. (See "How we think about charity" on our [research page](#)). On this claim, if we fund a project, we will be able to identify in advance where it lies on the impact curve; the impact curve of the organisation's own projects is not distributed according to a power law. Thus, *this* claim is different to the one about policy advocacy in the main text.

<sup>16</sup> Holden Karnofsky, 'Hits-Based Giving', Open Philanthropy Project, April 2016, <http://www.openphilanthropy.org/blog/hits-based-giving>; Susan Wolf Ditkoff and Abe Grindle, 'Audacious Philanthropy', *Harvard Business Review*, 1 September 2017, <https://hbr.org/2017/09/audacious-philanthropy>.

- The Rockefeller Foundation funded research into agricultural productivity, which is widely believed to have been a catalyst for the Green Revolution, saving over one billion people from starvation.
- Feminist philanthropist Katherine McCormick has been credited with being the sole funder of crucial early-stage development of the oral contraceptive pill.
- Professor Steve Teles has argued that the conservative legal movement profoundly changed the way the legal profession operates and the general intellectual stature of political conservatism.
- Philanthropically funded tobacco control advocacy helped to reduce smoking rates in the US from 42% in 1962 to 15% in 2016.
- In our climate change report, we showed that work by the Coalition for Rainforest Nations, funded primarily by the Linden Trust, played a major role in securing agreement on forestry at the Paris Agreement.

This suggests that policy philanthropy is like VC investing: it will be difficult to know in advance whether a particular project will succeed, but funding a portfolio of potential big wins means that your investment could nevertheless have high expected value.

### **Assessing track record in light of fat tails**

If policy advocacy does indeed follow a fat-tailed distribution, when we are evaluating the track record of policy organisations, we should focus on the big wins rather than the typical project. Consequently, we ask organisations to provide a few case studies detailing their most effective projects to date. We then estimate the benefits deriving from those case studies and the total costs the organisation has borne from the time of the oldest case study onwards. The reason we consider all costs from the oldest case study onwards, rather than simply the cost for each case study, is that—because of the fat-tailed distribution of impact—the successful projects are not representative of the overall past impact of an organisation.

### **Risk and picking winners**

The fat-tailed distribution of impact also means that it will be difficult to identify ahead of time which projects will succeed, whereas with direct interventions it is much easier to identify what the impact of a specific project will be. This means that supporting policy interventions is a high-risk/high-return investment: the typical investment is likely to fail, but the successful ones are likely to have a large impact.

Moreover, since we cannot identify winners in advance, if we are able to fill the funding gaps of multiple organisations, we should support a portfolio of organisations rather than picking only one. Since most of the expected value is driven by the bets that turn out to be big wins, we should only pick organisations that would be big wins if they did succeed: picking

organisations that produce only moderate benefits would not be optimal.<sup>17</sup> However, if a donor is not able to fill the funding gap of even one organisation, then donating to only one organisation might be the best approach.

### Patience and indirectness

The history of philanthropy suggests that not only is it difficult to identify in advance which organisation will have impact, advocacy often takes a long time to have an effect. For example, Teles and Schmitt discuss the advocacy effort required for Obamacare:

“Consider, for example, the campaign for US health care reform. The effort that culminated in 2010 was the result of decades of work, including a previous, high-profile failure in the early 1990s, waves of state-based reform, and numerous incremental efforts at the national level. Advocates invested hundreds of millions of dollars in initiatives ranging from media campaigns encouraging television producers to include stories of the uninsured, coalition-building projects, university- and think tank-based research, and grassroots initiatives. The basic outlines of reform policies were worked out well in advance, in advocacy groups and think tanks, which delivered a workable plan to presidential candidates. Important interest groups who could block reform, such as small business, had been part of foundation-supported roundtables seeking common ground for years. Technical problems had been worked out. And tens of millions of dollars had been set aside as long ago as 2007 for politically savvy grassroots advocacy initiatives targeted at key legislators. After a very long slog, the outcome was the Patient Protection and Affordable Care Act.”<sup>18</sup>

This suggests that the benefits of advocacy funding might not pay out until decades later. Having the patience to build a field around a particular problem is therefore an especially important niche for philanthropists to fill. Governments and the private sector do not usually have the long-term tolerance to be able to do this to the same extent, so this is a key comparative advantage of philanthropy.<sup>19</sup>

---

<sup>17</sup> Similarly, Peter Thiel argues that VCs should invest only in companies that could turn out to be unicorn companies. Blake Masters and Peter Thiel, *Zero to One: Notes on Start Ups, or How to Build the Future* (London: Virgin Books, 2015).

<sup>18</sup> Teles and Schmitt, “The Elusive Craft of Evaluating Advocacy.”

<sup>19</sup> Karnofsky, “Hits-Based Giving.”

## 5. Evaluating future impact

When you are deciding whether to donate to a charity, what matters is the effect that your donation has on the future impact of the charity. Three factors are central when evaluating future impact:

- Track record.
- Do they plan to work on a high-value project? How do they decide which projects to work on?
- Do their activities fit in with a plausible theory of change?

We discuss each of these in turn.

### Track record

When donating to charities implementing direct interventions, it is best to first identify a specific cost-effective intervention, and then find ways to support only that intervention. For example, our research partner GiveWell has identified deworming as a highly cost-effective programme and recommends donating to Sightsavers' deworming programme, rather than funding the organisation as a whole.<sup>20</sup> In this case, Sightsavers' track record provides a very strong indication of the quality of their work in the future because they will be implementing a similar programme.

In contrast, because the prospect of policy success is highly sensitive to context, effective charities involved in policy will have much more heterogeneous projects. For example, a climate change charity might advocate for a carbon price at one time and then for regulation of diesel engines at another, due to changes in the political context. Indeed, advocacy efforts often involve competition with a strategic adversary capable of adapting over time.<sup>21</sup> This means that an approach with a good track record might no longer be viable, and new adaptive approaches might not have a good track record. Therefore, a policy charity's track record is a fairly weak signal of its prospects for future success.

Nevertheless, track record is important in two respects: firstly it shows the ability of the organisation to get things done in a range of contexts; and secondly it shows the ability of the organisation to pick impactful projects.

An organisation could be good at achieving its aims without having had much impact in the past because it has previously focused on low-value projects. This might be a signal that the organisation will be impactful in the future because it has demonstrated the ability to achieve its aims. If the charity shifts to high-impact projects in the future, it will therefore stand a good chance of success. Thus, even though the past *impact* of the charity might be low, the charity

---

<sup>20</sup> GiveWell also recommends other deworming charities.

<sup>21</sup> Teles and Schmitt, "The Elusive Craft of Evaluating Advocacy."

might still be a good bet because its track record is a good indicator of the prospects of its future work.

However, past *impact*, as opposed to achieving aims, is an important indicator of an organisation's ability to select high-value projects.

### Project selection

Project selection is a key determinant of future impact. It is often useful to use proxies for impact to work out the value of a project that a charity plans to pursue. For example, if we were evaluating a climate charity, we might consider how much CO<sub>2</sub> is produced by the economic sectors or geographical areas the organisation wants to work in. If we were interested in assessing policy organisations working to promote evidence-based policy, we might consider the decision-making power of the policy-making institutions that these policy organisations are lobbying.

If a charity's policy selection process explicitly aims to choose cost-effective projects, then we can be more confident in the organisation's future cost-effectiveness. The key questions when assessing project selection are:

- Are they focused on maximising impact?
- Do they prioritise?
- Do they focus on outcomes rather than outputs?
- Do they make good use of research and evidence?

In policy advocacy, due to the variability of the political context, it will usually not be possible to assess all of an organisation's future activities. A recommendation of a policy organisation can to a large extent be seen as a bet on the organisation's team, ability to select high-value projects in the future, and strategic capacity to achieve its aims.<sup>22</sup>

### Theory of change

A theory of change defines the long-term goal of a campaign and works backwards to identify the most promising paths to success. A theory of change identifies necessary causal linkages in the path to a particular change and provides a rationale for why each stage in the causal process is necessary. They can be represented diagrammatically with boxes and arrows to signify different causal stages and the direction of influence. Theories of change are useful because, when planning, it is easy to neglect the ultimate goal of a project, the best way to achieve it, and potential barriers.<sup>23</sup>

It is not necessary for an organisation itself to have a theory of change if it is to be successful. An organisation might consistently choose effective ways forward without explicitly thinking

---

<sup>22</sup> Teles and Schmitt.

<sup>23</sup> This is similar to the idea of goal factoring. See alkjash, "Goal Factoring," LessWrong 2.0, accessed October 19, 2018, <https://www.lesswrong.com/posts/Cu5C5KhkoXhrPMLFN/goal-factoring>.

about projects in terms of a theory of change. However, constructing a theory of change may be useful for evaluators figuring out whether a particular campaign is likely to succeed or not.