

# Medium Investigation: Democratic and Institutional Reforms

## Table of Contents

### [Table of Contents](#)

#### [Introduction](#)

##### [Background and Motivation](#)

###### [Political short-termism](#)

##### [Scope and proposed reforms](#)

###### [General, democratic reforms](#)

###### [General, institutional reforms](#)

###### [Future-focused, democratic reforms](#)

###### [Future-focused, institutional reforms](#)

##### [Summary](#)

#### [Importance](#)

##### [Paths to impact and theory of change](#)

##### [Future-focused reforms](#)

###### [Existential risk mitigation](#)

###### [Imminent existential risk](#)

###### [Future existential risk](#)

###### [Potential paths to impact](#)

###### [Trajectory changes](#)

###### [Potential paths to impact](#)

##### [General reforms](#)

###### [Potential paths to impact](#)

##### [Summary](#)

#### [Neglectedness](#)

##### [Funding for general and future focused reforms](#)

###### [General reforms](#)

###### [Future-focused reforms](#)

##### [Summary](#)

#### [Tractability: will it get solved by default?](#)

##### [General considerations](#)

##### [Empirical data on the track record of reform efforts](#)

##### [Will it get solved by default?](#)

##### [Summary](#)

[Cost-effectiveness](#)

[What could a philanthropist do?](#)

[Future-focused reforms](#)

[Research](#)

[Advocacy](#)

[General reforms](#)

[Case for working on the problem](#)

[This cause seem highly important and neglected](#)

[Since few other longtermist funders are focused on this cause, the best opportunities are more likely to be under-funded](#)

[This cause is appealing to a variety of worldviews](#)

[Summary](#)

[Case against working on the problem & potential risks](#)

[Doubts about longterm impact](#)

[High intractability](#)

[Practical considerations for philanthropists](#)

[Summary](#)

[Potential risks for Founders Pledge](#)

[What are the main uncertainties in this analysis?](#)

[What should be the focus of future research?](#)

[Overall judgements](#)

[Key Sources](#)

## 1. Introduction

This medium investigation into democratic and institutional reform represents ~15 days of research. Medium investigations are the second stage of Founders Pledge's research process, and are meant to help us decide whether or not to produce a full, publishable Cause Area Report. To this end, medium investigations focus on how important a cause is, whether it is possible for a philanthropist to effect change, and whether a full investigation is likely to find cost-effective funding opportunities. They do not, however, include charity evaluations. They are also still limited in scope and depth.

We decided to do a medium investigation into democratic and institutional reforms because several effective altruists have suggested that this is a promising cause area and FP members have expressed interest in this cause.

## 1.1. Background and Motivation

Across their parliamentary, executive, and civil service departments, governments have a big influence in two main ways: by directing how public funds are spent, and by implementing policies that shape the actions of individuals and organisations within their jurisdiction. For example, the US Federal Government alone [spends about 20% of the US's GDP each year](#), or over \$4 trillion.<sup>1</sup> How this funding is allocated depends on who holds power, and the structure of the democratic system affects who wields power. As a simple example: in the US, positions like Secretary of State are appointed by the president. In the UK, the equivalent position (Foreign Secretary) is chosen from the set of elected MPs (a much smaller group of candidates). This kind of institutional variation meaningfully affects the options available to and selected by government decision makers.<sup>2</sup>

Some of this spending and many of these decisions have large plausible impacts on the longterm future. For example, governments can fund existential risk research and initiatives, and make many decisions that affect how technologies are governed or the likelihood of global catastrophes like war. This means it's reasonable to think that different democratic or institutional systems than our current ones could lead to better outcomes. For example:

- Voting systems other than plurality (or first-past-the-post) voting could lead to outcomes that satisfy more voters, reduce political polarization, or avoid vote splitting
- Efforts to represent the needs/preferences of future generations could lead to better long-term policy outcomes, if these preferences are currently underweighted because future generations don't vote in our system
- Some people have suggested that voters, politicians, and policymakers are biased or incentivized towards considering the short-term effects of their actions and neglecting the long-term effects. Initiatives to raise awareness, provide information, or mandate the consideration of long-term effects may be able to mitigate this

---

<sup>1</sup> Of course, not all of this money is spent on issues that are highly relevant to longtermists. Roughly  $\frac{2}{3}$  of this is spent on mandatory programs such as medicare, medicaid, and social security that seem more difficult to influence and less relevant from a longtermist point of view. But  $\frac{1}{3}$  of this is discretionary spending that must be appropriated by Congress each year. In 2019, ~half of this was Defense spending and ~half was split among issues like transportation, education, housing, science, etc.

Let's say  $\frac{1}{4}$  of discretionary spending impacts issues relevant to longtermism - that leaves ~\$160B per year of important spending in the US alone

<sup>2</sup> We have about 60% confidence in the claim that institutional structures are a major determinant of the variation in outcomes due to policymaking. This is still speculative, and we would be interested in seeing analyses or arguments that contend that the institutional structure of government is actually not correlated with different outcomes. For example, we can imagine worlds where variables like education levels or demographics have a larger influence on the quality of government decisionmaking

Working towards the implementation of these systems could be highly-effective, not least because this work seems likely to be highly leveraged. If successful, it could affect much larger amounts of money than is spent on advocacy, and if sustained, these reforms could have effects over long time periods by changing the trajectory of the long-term future.

## Political short-termism

A major longtermist motivator of democratic and institutional reforms is mitigating political short-termism. Governments are run by presently existing people and, in democracies, are elected by and accountable to presently existing people. In particular, elected officials work in relatively short election cycles, lasting just a few years at most. As a result, it is plausible that many policy and decision makers are concerned primarily with the impacts of their actions on the scale of just a few years. [John and MacAskill \(ms\)](#) outlines a variety of causes for political short-termism:<sup>3</sup>

- Epistemic determinants – features of political actors' state of knowledge that prevent (even properly-motivated) actors from adopting appropriately long-termist policy
  - Rational discounting of future impacts due to a lack of information about the future
  - Irrational discounting of future impacts due to cognitive biases and attentional asymmetries between the future and the nearby past
- Motivational determinants – features of political actors' goals and motivations that lead (even well-informed) actors to wrongfully discount the future
  - positive rate of pure time preference
  - Self-interest
  - relational partiality: political actors are sometimes motivated to benefit those that they have close relations with, such as friends family or their community, at the expense of the rest of the current and future generations
  - cognitive biases
- Institutional determinants – features of political actors' institutional context that strip the political means from (even well-informed, properly-motivated) actors who could otherwise adopt more appropriately long-termist policy, or which make political actors less well-informed or less well-motivated
  - election incentives to prioritise policy with near-term results and visible benefits for which politicians and parties can take credit
  - economic dependency on various firms and other bodies that exert pressure on political actors to use short auditing durations, e.g. through performance indicators with short-term goals and positive discount rates, inadequate credit-tracking over longer time-frames, and budget windows with short time-frames all incentivize political leaders to shift benefits to the short-term and costs into the future.

---

<sup>3</sup> [Mindmap](#) linking proposals to root causes.

- short media cycle requires political actors to react and respond swiftly to political issues
- political polarisation detracts from careful, collective deliberation
- omnibus bills are passed or rejected before they can be carefully discussed in full
- lack of strong commitment devices to ensure that governments act on past promises leads to low levels of trust in long-term policy proposals
- voters and elites oppose future-beneficial policy promises which might be reneged due to low levels of trust
- institutions may be too weak to reliably bring about long-run outcomes or they may be plagued by collective action problems that undermine successful coordination

Political short-termism plausibly leads to policy and decision making that is far from optimal for future generations. This suggests a high importance of reforms that target these root causes. While the above considerations seem plausible to us, we have not investigated these claims deeply beyond the John and Macaskill paper cited above. We are still uncertain about the extent to which the above factors drive political short-termism, how much political short-termism reduces the value of the far-future, and how mitigating the above factors would change policy outcomes.

## 1.2. Scope and proposed reforms

In this report, we consider both democratic and institutional reforms. We define these reforms as follows:

- A **democratic reform** is an intervention that changes the way representative citizens decide how governments are formed or influenced. Interventions in this category include alternative voting systems and campaign finance rules
- An **institutional reform** is an intervention that systematically changes the way public bodies function. Interventions in this category include creating new positions or bodies within policymaking institutions.

For this investigation, we are interested primarily in the effect of these reforms from a *longtermist* point of view. The main outcome we consider is total well-being. We may also discuss the effect of reforms on a wider range of values, such as equality, justice and democracy.

The terms “democratic” and “institutional” are broad. For practical purposes, we exclude some interventions that could arguably fall within these categories. Some notable interventions outside the scope of this document include peace-building, promoting democracy internationally, and protecting human rights.

Since we are primarily focused on the effects of these reforms from a longtermist point-of-view, we also do not closely examine some popular reforms that lack a plausible longterm path to impact.

We categorize reforms in two dimensions. First, we distinguish democratic reforms from institutional reforms, as defined earlier. Second, we distinguish future-focused reforms from general reforms:

- **General** reforms have broad effects. While they *may* affect the long-term future, this is not their primary goal
- **Future-focused** reforms seek to change political processes with the specific goal of improving the long-term future (e.g. legislative house for future generations)

Here we briefly describe the following reforms:

Possible reforms	<i>Democratic</i>	<i>Institutional</i>
<i>General</i>	<ul style="list-style-type: none"> <li>• Alternative voting systems</li> <li>• Campaign finance</li> <li>• Citizens' assemblies and sortition</li> </ul>	<ul style="list-style-type: none"> <li>• Improving decision-making by implementing prediction markets or forecasting techniques</li> </ul>
<i>Future-focused</i>	<ul style="list-style-type: none"> <li>• Voting systems that give more weight to voters who may have longer time horizons</li> <li>• Creating positions or bodies that inform policymakers and the public about issues that affect future generations</li> <li>• Mandates that require policymakers to consider the effect of legislation on future generations</li> <li>• Creating positions or bodies that explicitly represent the interests of future generations</li> <li>• Research to generate new proposals or analyze existing ones</li> </ul>	<ul style="list-style-type: none"> <li>• Supporting non-partisan bodies focused on future generations</li> <li>• Research to generate new proposals or analyze existing ones</li> </ul>

## General, democratic reforms

### *Voting systems*

- Many democratic reform proposals focus on changing the way people vote to elect governments

- There's a large literature on voting theory, analyzing the strengths and weaknesses of different voting systems.
- Notable proposals include switching to approval voting, alternative/runoff voting, proportional representation, quadratic voting.
- Since many electoral reform proposals focus on changing the voting system, this issue seems harder for a marginal funder to affect
  - Voting system reforms usually have to go to referenda or ballot initiatives which fail more often than not
- The relative merits of different systems are contentious
  - First past the post is generally considered to be particularly bad. Given this, Pablo Stafforini writes "I think the priority for EAs is to support whichever alternatives to plurality voting are most viable in a particular jurisdiction, rather than obsess over which of these alternatives to plurality is the absolute best."
  - Jameson Quinn cautions against focusing too much on sets of criteria, as Kenneth Arrow did, since it's impossible for any given system to fulfill all of them. Instead, think more in terms of weights on different criteria to compare systems.<sup>4</sup>
- Voters are often poorly informed and so it's unlikely that more representative systems are strictly better systems (notwithstanding democracy as intrinsically valuable)
- Examples of initiatives in this space:
  - [FairVote](#) advocates for the use of instant runoff voting in elections at various levels in the US
  - The [Center for Election Science](#) advocates for the use of approval voting at various levels in the US
  - There are several other organisations that run campaigns and initiatives in this space

### *Campaign finance*

- Campaign finance reform is another popular general proposal, [especially in the US](#)
- Usually the proposal is to limit the amount of money wealthy private donors can give to political parties, increase transparency, or increase public funding to political parties to level the playing field
- This doesn't seem that promising of an area to me because studies of the influence of money on electoral outcomes generally find that campaign finance has a causal effect only in specific circumstances<sup>5</sup>

---

<sup>4</sup> "Since it's impossible to meet all desirable criteria in all cases, I'd rather look at things in a more probabilistic and quantitative way: how often and how badly does a given system fail desirable criteria" (<https://www.lesswrong.com/posts/D6trAzh6DApKPhbv4/a-voting-theory-primer-for-rationalists>)

<sup>5</sup> "But decades of research suggest that money probably isn't the deciding factor in who wins a general election, and especially not for incumbents. Most of the research on this was done in the last century, Bonica told me, and it generally found that spending didn't affect wins for incumbents and that the impact for challengers was unclear. Even the studies that showed spending having the biggest effect, like one that found a more than 6 percent increase in vote share for incumbents, didn't demonstrate that money causes wins. In fact, Bonica said, those gains from spending likely translate to less of an advantage today, in a time period where voters are more stridently partisan. There are probably fewer and fewer people who are going to vote a split ticket because they liked your ad" (Maggie Koerth, "How Money

- [Our impression: 90% confidence this isn't the best intervention, 60% confidence campaign finance doesn't have a big effect on electoral outcomes]
- [FollowTheMoney.org has a list of organisations](#) that advocate for reforms of laws around campaign finance in the US. The relevance/activeness of many of these is questionable, but the list includes the [American Civil Liberties Union](#), which advocates for public funding of political campaigns. Reclaim the American Dream has a [similar list](#), which largely overlaps

### *Citizens' assemblies and sortition*

- Citizens' assemblies are deliberative bodies of randomly selected citizens that provide non-binding advice to the government
- Sortition is the practice of choosing officials by lottery rather than election
- Engage with subject matter experts
- Discussion among diverse people and points of view
- Ideally, the resulting advice and decisions come from a better informed but representative citizenry
- Non-exhaustive list of relevant organisations:
  - [Citizens' Assembly](#)
    - Information on citizens' assemblies in the UK
  - [Sortition Foundation](#)
    - Promotes the use of sortition (lotteries) rather than or as well as elections globally
  - [G1000UK](#)
    - Advocates for the use of sortition at the community level in the UK

## General, institutional reforms

### *Improving government decision-making*

- As an example of how this could be advanced, there was movement recently to get more superforecasters into government advisory positions in the UK<sup>6</sup>
- For example, the Institute for Government has published [at least one report](#) making the case for institutional reform of government in the UK

## Future-focused, democratic reforms

### *Voting systems that give more weight to voters with longer time-horizons*

Affects Elections", 2018, *FiveThirtyEight*,

<https://fivethirtyeight.com/features/money-and-elections-a-complicated-love-story/>)

<sup>6</sup> Though this proved controversial due to the association with Dominic Cummings: "[Superforecasting] could be useful in areas from finance, to charities working out how they should distribute aid. And CIA analysts wrote a paper calling for the US intelligence service to look for the characteristics of superforecasters when recruiting, rather than prioritising applicants' grades. Prof Tetlock told the BBC he did not think superforecasting should be linked to a particular political point of view. He said most people would want their leaders to be "informed by the most accurate possible estimates of the consequences of the options on the table"" (BBC News, "Andrew Sabisky: What is superforecasting?", 2020, <https://www.bbc.co.uk/news/uk-politics-51545541>)

- Working under the assumption that political short-termism is rooted in voter preferences, and that younger voters have longer time-horizons, Will Macaskill has floated the idea of [age-weighted voting](#). This system would give votes by younger people more weight
- But the idea has not been scrutinized thoroughly. Macaskill writes that the proposal “would need a lot more investigation before [he’d] want to endorse it”
- I don’t think there’s any concrete funding opportunity working on this idea at this time

#### *Bodies that explicitly represent the interests of future generations*

- Futures assemblies: deliberative bodies of randomly selected citizens to provide non-binding advice to the government, with an explicit mandate to represent the interests of future generations
- Legislative houses for future generations: an upper legislative house tasked with attending to the interests of future generations that must pass legislation in order for it to become law

### Future-focused, institutional reforms

#### *Initiatives to increase the consideration of future generations by policymakers*

- These include efforts to bring more analysis of the effects of policies on future generations to the attention of policymakers
- Examples include:
  - [All-Party Parliamentary Group for Future Generations](#)
  - Creating an [Ombudsman for Future Generations](#)
  - Direct advocacy, e.g. [Alpenglow](#)
  - [Methodology to assess public policies from the perspective of intergenerational fairness](#), supported by the [Gulbenkian Foundation](#) and the [School of International Futures](#)
- According to Jones et al., six countries currently have or had official institutions for future generations within the government:

Country	Dates of operation	Position with respect to executive and legislature	Scope	Functions and Powers
Finland	1993-	Standing Committee of Parliament	Futures in general; can choose own scope	Research/advisory Education
Hungary	2008-2012	Structurally independent from government	Issues which may affect the constitutional right to a healthy environment	Research/advisory Complaints investigation Legal enforcement
Israel	2001-2006	Parliamentary committee	Environment, natural resources, science, development, education, health, state economy, demography, planning	Research/advisory Initiate legislation Veto legislation

			and building, quality of life, technology, law, any other matter considered relevant	
Scotland	2005-	Structurally independent from government	Futures in general; can choose own scope	Research/advisory Education
Singapore	2009-	Within the Prime Minister's Office	Risk and futures; can choose own scope	Research/advisory Education
Wales	2016-	Structurally independent from government	Sustainable development	Research/advisory Recommendations are binding

Source: [Jones et al., 2018, p. 4](#)

Jones's work is missing at least one relevant body: the UK's All-Party Parliamentary Group for Future Generations.

### 1.3. Summary

- We're investigating this cause because some researchers have suggested that political short-termism is likely to cause policymakers to make decisions that harm future generations, and that democratic or institutional reforms can alleviate this problem
- It is also of interest to FP members
- We investigate reforms in two dimensions: first, we look at both institutional and democratic reforms. Second, we look at both general and future-focused reforms
- We have found a variety of proposals in each of these categories
- In the following sections, we consider the importance, neglectedness, and tractability of this cause

## 2. Importance

Naively, democratic and institutional reforms look very important: billions of present and future people are, or will be, affected by democracies (either directly by living in democracies or indirectly through the foreign policy of democracies), so democratic and institutional reforms could, in principle, affect a very large number of people. In assessing the importance of work in this area, it is helpful to think more concretely and specifically about how reforms could have a long-term impact.

When evaluating a problem, we usually think of its *importance* in terms of how many people/beings it affects and the extent to which it affects their well-being.<sup>7</sup> In this sense, importance is a measure of the potential value at stake: how much we would improve the world if we completely solved the problem. In this context, we're not so much evaluating a problem, but rather the opportunity to influence the future through democratic and institutional reforms

<sup>7</sup> Of course, one can also consider importance in terms of other, non-utilitarian values.

(which might solve some problems and/or bring about additional benefits). We can think of importance here in a similar way though: how large the benefits from democratic and institutional reforms could be, mainly setting aside questions of tractability and counterfactual impact for now. We'll first introduce some concepts and terminology that help think about the importance, before offering initial (and necessarily speculative) thoughts on the importance of democratic and institutional reforms.

## 2.1. Paths to impact and theory of change

Call a "path to impact" (or sometimes just "path" for short) a causal chain from a specific reform or a broad category of reforms to a specific policy or decision making improvement to a specific better outcome. For example, a futures assembly could stress that we aren't taking sufficient action against climate change, resulting in improved climate policy, and consequently less severe climate change, more global stability in international relations, less existential risk and greater prosperity. This is an illustrative example of a path to impact. We are interested here in getting a sense of how large the impact could be along various paths to impact as well as how broad the class of paths to impact could be.

The path to impact of most democratic and institutional reforms assumes the following simple theory of change:

Reform → Improved policy and/or decision making → Better outcomes
---

There are three main categories of improved long-term outcomes due to direct interventions:<sup>8</sup>

- Lower existential risk
- Positive trajectory change
- Speeding up positive developments

Lowering existential risk is reducing the chance that an existential catastrophe occurs. An existential catastrophe is an event, such as human extinction, that would permanently and drastically curtail humanity's potential.<sup>9</sup>

To understand what a trajectory change is, it will help to dig a little deeper into the details of trajectory changes in Nick Beckstead's Dissertation. Beckstead distinguishes between historically contingent events and non-contingent events.<sup>10</sup> Non-contingent events are events

---

<sup>8</sup> 'Beckstead, Nick--On the Overwhelming Importance of Shaping the Far Future (Better Formatting).Pdf', Google Docs, accessed 16 December 2020, [https://drive.google.com/file/d/0B8P94pg6WYC1c0XSUVYS1BnMkE/view?usp=embed\\_facebook](https://drive.google.com/file/d/0B8P94pg6WYC1c0XSUVYS1BnMkE/view?usp=embed_facebook).

Beckstead also considers speeding up development as a benefit. We set this aside for now.

<sup>9</sup> 'Existential Risks: Analyzing Human Extinction Scenarios', accessed 8 January 2021, <https://www.nickbostrom.com/existential/risks.html>.

<sup>10</sup> "If Thomas Edison had not invented the light bulb, someone else would have done it later. In this sense, it is not historically contingent that we have light bulbs, and the most obvious benefits from Thomas Edison inventing the light bulb are proximate benefits and benefits from speeding

that counterfactually would have happened anyway, though probably later than they did in fact happen. The canonical example is technological development, such as the invention of the light bulb: if Thomas Edison had not invented it, then someone else would have done so a short while later. Historically contingent events are events that counterfactually would not have happened or would have happened very differently. Historically contingent events with far-reaching consequences can have enormous effects on the trajectory of civilisation. Beckstead suggests that the rise of Christianity, the creation of the US Constitution, and the influence of Marxism are examples of historically contingent events.<sup>11</sup> Historically contingent events are trajectory changes.

Beckstead also distinguishes between broad and targeted ways of shaping the future:<sup>12</sup>

The general distinction is that broad approaches focus on unforeseeable benefits from ripple effects, whereas targeted approaches aim for more specific effects on the far future, or aim at a relatively narrow class of positive ripple effects.

A broad approach could be to improve education, or democratic institutions, with the aim of having many positive, but unforeseeable, effects. Improving biosecurity policy is a targeted approach as it primarily aims to reduce biological global catastrophic risk specifically. In the terminology of this investigation, broad approaches to shaping the future have very many paths to impact, many of which are likely unforeseeable, whereas targeted approaches aim at specific targeted paths to impact. Of course, there is a spectrum of ways of shaping the future between broad and targeted extremes.

It can be helpful to distinguish clearly between existential risk mitigation and trajectory changes. Avoiding an existential catastrophe is a kind of trajectory change: "If we ever prevent an

---

up development. Something analogous is probably true of many other technological innovations: computers, candles, wheelbarrows, object-oriented programming, and the printing press, to give an arbitrary list of examples." 'Beckstead, Nick--On the Overwhelming Importance of Shaping the Far Future (Better Formatting).Pdf', Google Docs, accessed 16 December 2020, [https://drive.google.com/file/d/0B8P94pg6WYC1c0IXSUVYS1BnMkE/view?usp=embed\\_facebook](https://drive.google.com/file/d/0B8P94pg6WYC1c0IXSUVYS1BnMkE/view?usp=embed_facebook). Pg. 7.

<sup>11</sup> "There have been other events that were historically contingent, and changed the course of history significantly. Potential examples include: the rise of Christianity, the creation of the US Constitution, and the influence of Marxism. Various aspects of Christian morality influence the world today in significant ways, but the fact that those aspects of morality, in exactly those ways, were part of a dominant world religion was historically contingent. And therefore events like Jesus's death and Paul writing his epistles are examples of trajectory changes. Likewise, the US Constitution was the product of deliberation among a specific set of men, the document affects government policy today and will affect it for the foreseeable future, but it could easily have been a different document. And now that the document exists in its specific legal and historical context, it is challenging to make changes to it, so the change is somewhat self-reinforcing." 'Beckstead, Nick--On the Overwhelming Importance of Shaping the Far Future (Better Formatting).Pdf', Google Docs, accessed 16 December 2020, [https://drive.google.com/file/d/0B8P94pg6WYC1c0IXSUVYS1BnMkE/view?usp=embed\\_facebook](https://drive.google.com/file/d/0B8P94pg6WYC1c0IXSUVYS1BnMkE/view?usp=embed_facebook). Pgs. 7-8.

<sup>12</sup> 'Beckstead, Nick--On the Overwhelming Importance of Shaping the Far Future (Better Formatting).Pdf', Google Docs, accessed 16 December 2020, [https://drive.google.com/file/d/0B8P94pg6WYC1c0IXSUVYS1BnMkE/view?usp=embed\\_facebook](https://drive.google.com/file/d/0B8P94pg6WYC1c0IXSUVYS1BnMkE/view?usp=embed_facebook). Pg. 10.

existential catastrophe, that would be an extreme example of a trajectory change.”<sup>13</sup> Beckstead uses “trajectory change” to refer to other, smaller changes to the future trajectory. Similarly, we will use “trajectory change” as a shorthand for “non-existential catastrophe trajectory changes”. Averting an existential catastrophe could make the difference between a highly prosperous future and an empty, or drastically curtailed, future. Causing a trajectory change could make the difference between a somewhat good and an excellent future.

We will largely leave aside the third class of benefits, speeding up development, in this investigation as these will often be less valuable than the first two and because democratic and institutional reforms seem better placed to bring about the first two.

Along some paths to impact, the interests of the current and future generations are closely aligned (e.g. mitigating imminent existential risks), whereas along others, they compete (e.g. mitigating climate change). When intergenerational interests compete, usually the world would be better if the current generation acted against its own interests, in order to benefit future generations, as the aggregate of all future generations is so much larger in expectation.<sup>14</sup> But only the current generation can decide how to act and we should expect the current generation to act mainly in its own interests, unless there are mechanisms compelling it to act otherwise. Future-focused democratic and institutional reforms could provide such mechanisms. This can be a useful lens for viewing paths to impact and considering how much value they could bring about.

## 2.2. Future-focused reforms

### Existential risk mitigation

Democratic and institutional reforms could indirectly mitigate existential risk. For example, such reforms may improve global coordination, improve the ability of government bodies to solve problems and manage risks across a broad range of domains, or facilitate better policy making surrounding specific risks (e.g. AI, biosecurity). In this sense, democratic and institutional reforms could be an existential security factor, broadly safeguarding the future. We currently think it’s likely that direct approaches to existential risk mitigation that directly target specific risks are more effective, but we haven’t modelled this in detail.<sup>15</sup>

---

<sup>13</sup> ‘Beckstead, Nick--On the Overwhelming Importance of Shaping the Far Future (Better Formatting).Pdf’, Google Docs, accessed 16 December 2020, [https://drive.google.com/file/d/0B8P94pg6WYClc0lXSUVYS1BnMkE/view?usp=embed\\_facebook](https://drive.google.com/file/d/0B8P94pg6WYClc0lXSUVYS1BnMkE/view?usp=embed_facebook). Pg. 7.

<sup>14</sup> It is often easiest to think in binary terms: interests compete or are aligned. But in reality, this is of course a spectrum and we should bear that in mind.

<sup>15</sup> On existential risk and security factors, Ord writes “just because existential risk declines as some other goal is pursued doesn’t mean that the other goal is the most effective way to secure our future. Indeed, if the other goal is commonsensically important there is a good chance it is already receiving far more resources than are devoted to direct work on existential risk [...] I think it is likely that there will only be a handful of [such factors] that really compete with the most important existential risks” (The Precipice, p. 180)

It might be helpful to distinguish between imminent existential risk and future existential risk, as the potential value of future-focused reforms looks different in each case, and therefore depends on how urgent we think existential risks are.

## Imminent existential risk

In the case of mitigating imminent existential risk, intergenerational interests are closely aligned as all generations want to survive (assuming they have positive lives). We might naively expect future-focused reforms to be of little value in areas in which intergenerational interests are closely aligned as the current generation will self-interestedly act in accordance with the interests of future generations anyway. It is not clear that this is the case though. Even though each generation has a strong interest in surviving and in allocating resources to mitigating existential risks, the aggregate interests of all future generations are likely to be much stronger than the interests of the current generation. As a result, the current generation is likely to underinvest in existential risk mitigation, from an impartial intergenerational perspective. This is a kind of scope neglect. It can also be viewed as a negative externality: if we go extinct, the current generation does not bear the full cost of extinction.

A more concrete, down-to-earth example: suppose you are deciding how often to go skydiving. You will enjoy the thrill of skydiving but you do not want to die, and you have to weigh these interests. You want to go skydiving frequently enough that you get a good thrill but not so frequently that you expose yourself to a high risk of death.

Now suppose you are again deciding how often to go skydiving, but every time you go skydiving, you have to skydive with all your children and grandchildren tied to you. Furthermore, your children and grandchildren do not get to choose whether they skydive: if you do, then so do they. In this case, the downside risk is greater, so you would probably decide to go skydiving less than you would in the first case.

*Hypothesis:* Existential risk mitigation more closely resembles the second case but policy and decision makers (and people more broadly) usually act more in line with the first case. These examples do not offer a perfect analogy but they demonstrate how the current generation could be underinvesting in existential risk mitigation (i.e. skydiving too frequently) even though every generation has strong interests in preventing existential risks.

Overall, we are unsure of how important this consideration is and think the naive view given at the start of this subsection should carry at least non-negligible weight (30% confidence in the naive view, not very resilient). It could be the case that in practice, the current generation acts close-to-optimally and that marginal improvements would be of little value. However, if the future holds very many future generations, then even very small reductions in existential risk could be highly valuable and so marginal improvements in policy and decision making could still be highly valuable. A similar scope-neglect-based argument could be made about marginally better trajectory changes, given aligned intergenerational interests.

## Future existential risk

In the case of mitigating future existential risk (e.g. many people think that AI existential risk is very low in the next decade or so but could be much higher in the future), then intergenerational interests compete. Generations in the time of high existential risk and beyond have strong interests in taking pre-emptive action to mitigate existential risk that compete with earlier generations' interests to improve their own lives rather than prepare for future periods of high existential risk. As a result, there is a stronger case for the importance future-focused reforms if existential risk mitigation is less urgent.

### Potential paths to impact<sup>16</sup>

- Future-focused research body informs policy makers and the public of risks posed by emerging technology (e.g. AI, biorisk) and raises the public profile of such risks
  - policy makers are better informed for mitigating risks
  - + public is more concerned about these risks, discount the future less, and votes for officials who prioritise these risks more
    - improved policy surrounding emerging technology, stronger incentives to regulate and cooperate
    - reduced existential risk
- Posterior impact assessments require policy makers to consider the effects of legislation on future generations
  - greater scrutiny of and consideration of the effects of legislation surrounding emerging technologies
    - improved policy surrounding emerging technology
    - reduced existential risk
- Futures assembly mandated to represent the interests of future generations
  - improved quality of discussion surrounding the effects of emerging technologies on future generations
    - stronger incentives for policy makers to take interests of future generations into account in policy making surrounding emerging technologies
      - improved policy surrounding emerging technology
      - reduced existential risk
- Legislative house for future generations
  - increased scrutiny of legislation, taking the interests of future generations into account to a greater extent
    - improved policy surrounding emerging technology
    - reduced existential risk

## Trajectory changes

It is more difficult to point to specific positive trajectory changes that reforms could bring about as the causal pathways will likely be even broader and less direct. The class of potential

---

<sup>16</sup> These examples are illustrative and speculative.

trajectory changes is much larger than the class of existential risks and many trajectory changes are difficult to predict in advance.<sup>17</sup> Beckstead writes that “No one could have predicted the persistent ripple effects that Jesus’s life had, for example.” Beckstead also notes that in some cases, such as the writing of the US Constitution, it is clear that an event could turn out to be a trajectory change but very difficult to specify beforehand what we should concretely do. As a result, he thinks that promising trajectory change attempts may be very broad and indirect, such as improving education, improving science, or improving political systems:<sup>18</sup>

Though it seems unlikely that the far future will inherit many of our institutions exactly as they are, it is not hard to believe that various aspects of the far future—including social norms, values, political systems, and technologies—will be path dependent on what happens now, and often in a suboptimal way. In general, it is reasonable to assume that if there is some problem that might exist in the future and we can do something to fix it now, future people would also be able to solve that problem. But if values or social norms change, they might not agree that some things we think are problems really are problems. Or, if a certain standards or conventions get sufficiently entrenched, some problems may be too expensive to be worth fixing.<sup>19</sup>

More specifically, making decision and policy making processes more longtermist (thinking about the very long-term consequences of their decisions and policies to a greater extent), could be a valuable trajectory change, potentially leading to a host of further trajectory changes.

Future-focused reforms are, perhaps, unusually foreseeably valuable, compared to other trajectory changes. It seems straightforwardly very good to have policy and decision makers thinking more about the effects of their policies and decisions on future generations. Furthermore, democratic and institutional reforms directly impact policy and decision making, which directly impact lots of people to a great extent. Compare to improving education: this also seems straightforwardly good but it is much further removed from large levers that affect lots of people to a great extent. In this way, future-focused reforms are targeted relative to other possible trajectory change interventions, though still much broader than other longtermist interventions (e.g. improving biosecurity policy), and more foreseeably good than other trajectory change interventions.

---

<sup>17</sup> ‘Beckstead, Nick--On the Overwhelming Importance of Shaping the Far Future (Better Formatting).Pdf’, Google Docs, accessed 16 December 2020, [https://drive.google.com/file/d/0B8P94pg6WYClc0IXSUVYS1BnMkE/view?usp=embed\\_facebook](https://drive.google.com/file/d/0B8P94pg6WYClc0IXSUVYS1BnMkE/view?usp=embed_facebook). Pg. 8.

<sup>18</sup> “In other cases—such as the framing of the US Constitution—it’s clear that a decision has trajectory change potential, but it would be hard to specify, in advance, which concrete measures should be taken. Because of this, promising ways to create positive trajectory changes in the world may be highly indirect. Improving education, improving our children’s moral upbringing, improving science, improving our political system, spreading humanitarian values, or otherwise improving our collective wisdom as stewards of the future could create many small, unpredictable positive trajectory changes.” ‘Beckstead, Nick--On the Overwhelming Importance of Shaping the Far Future (Better Formatting).Pdf’, Google Docs, accessed 16 December 2020,

[https://drive.google.com/file/d/0B8P94pg6WYClc0IXSUVYS1BnMkE/view?usp=embed\\_facebook](https://drive.google.com/file/d/0B8P94pg6WYClc0IXSUVYS1BnMkE/view?usp=embed_facebook). Pg. 8.

<sup>19</sup> ‘Beckstead, Nick--On the Overwhelming Importance of Shaping the Far Future (Better Formatting).Pdf’, Google Docs, accessed 16 December 2020, [https://drive.google.com/file/d/0B8P94pg6WYClc0IXSUVYS1BnMkE/view?usp=embed\\_facebook](https://drive.google.com/file/d/0B8P94pg6WYClc0IXSUVYS1BnMkE/view?usp=embed_facebook). Pg. 8.

If the trajectory of civilisation is very contingent on the current state, then there could be important path dependencies such that the value of the future could vary significantly depending on how our institutions work. These could be broadly due to political norms, institutions and/or in specific policy areas affected by reforms. As a result, the value of a reform from a longtermist perspective depends on the background rate of trajectory-altering decisions being taken. If such decisions come about once every 100 years (e.g. framing a constitution), then implementing a democratic reform a couple of years earlier does not seem that valuable since it's unlikely that an trajectory-altering decision will come about in those years.

We should also bear in mind the possibility of [avoiding astronomical suffering](#)<sup>20</sup> (s-risks), a type of trajectory change<sup>21</sup> that could be even more important than avoiding human extinction. We haven't considered this possibility in much detail but note that strong democratic institutions seem good for avoiding very bad outcomes, e.g. through strong checks and balances, or by placing more power with the informed median citizen (who would object to some very harmful practices and elect less harmful leaders).

### Potential paths to impact<sup>22</sup>

- Voting system that gives greater weight to younger/future generations
  - increased electoral weight given to voters with longer to live
  - stronger incentives for elected officials to take long-term effects of policies and decisions into account
  - broadly improved policy and decision making
  - improved trajectory
- Posterior impact assessments require policy makers to consider the effects of legislation on future generations
  - greater scrutiny of and consideration of the effects of legislation with long-term impacts
  - broadly improved policy with long-term impacts
  - improved trajectory
- Futures assembly mandated to represent the interests of future generations
  - improved quality of discussion surrounding the long-term effects of policies
  - stronger incentives for policy makers to take interests of future generations into account in policy making
  - broadly improved policy with long-term impacts
  - improved trajectory

## 2.3. General reforms

General reforms seem most valuable along paths for which intergenerational interests are well-aligned. In this context, general reforms can empower the current generation to make

---

<sup>20</sup> Brian Tomasik, 'Risks of Astronomical Future Suffering', *Center on Long-Term Risk* (blog), 10 April 2015, <https://longtermrisk.org/risks-of-astronomical-future-suffering/>.

<sup>21</sup> Note that s-risks are sometimes categorised as existential risks.

<sup>22</sup> These examples are illustrative and speculative.

better decisions and policies for themselves that also benefit future generations, for example mitigating urgent existential risk. Additionally, some general reforms could help mitigate future existential risk by putting future decision and policy makers in stronger positions for mitigating existential risk in their time. However, Ord argues we should usually expect to be able to mitigate existential risk better by targeting risks directly rather than through general reforms.

Some general reforms could mitigate political short-termism by combatting some of the root causes of political short-termism identified by John and MacAskill, e.g. by improving the quality of political debate, reducing political polarisation, improving the quality of available information.

The most valuable trajectory changes that reforms could bring about may well be increased concern for future generations (or changes that stem from that) that future-focused reforms (and much less so general reforms) could bring about.

The overall value of a reform is (roughly<sup>23</sup>) the sum of the value along all its paths to impact, which can be decomposed as the average value of the reform's path to impact multiplied by the number of paths to impact. The average value of future-focused reforms' paths to impact seem larger, potentially *much* larger, than the average value of general reforms' paths to impact. But future-focused reforms do not lose much breadth by targeting benefits for future generations. They may even be broader: though the benefits of future-focused reforms are targeted in the future, there could be a larger number of paths to impact than from general reforms. As a result, future-focused reforms seem likely to be more valuable than general reforms.

#### Potential paths to impact<sup>24</sup>

- Improved voting system
  - candidates need wider support in order to be elected
    - policy makers are incentivised to enact policy more in line with a larger proportion of the citizenry
    - broadly improved policy
    - improved trajectory
- Campaign finance reform
  - the wealthy have relatively less power and the median citizen has relatively more power
    - policy makers are incentivised to enact policy more in line with a larger proportion of the citizenry
    - broadly improved policy
    - improved trajectory
- Citizens' assembly
  - improved quality of discussion across various policy areas
    - stronger incentives for policy makers to take interests of the informed citizenry into account

---

<sup>23</sup> Some care is needed to avoid double-counting impact along overlapping paths to impact.

<sup>24</sup> These examples are illustrative and speculative.

- broadly improved policy
  - improved trajectory
- Citizens' assembly
  - improved quality of discussion across diverse range of people, communicated to wider public
    - reduced political polarisation
      - reinforces constructive, high-quality debate
        - broadly improved policy
          - improved trajectory

## 2.4. Summary

- Naively, we should expect democratic and institutional reforms to be very important given the scope of influence government has
- But for these reforms to be important from a longtermist point of view, we need to consider exactly how they affect the long-term future:
  - They could reduce existential risk
  - They could cause a trajectory change
- Reducing existential risk: this seems plausible through a few mechanisms (reducing the chance of war or improving governance of emerging technologies, for example) but Ord suggests we should have a skeptical prior about the value of broad reforms vs. working on those risks directly. We have not seen a strong case for any general reform significantly reducing existential risk in a way that seems comparable to direct work, though we have not spent much time looking outside of the EA literature.
- Causing a trajectory change: the importance here is a function of three main variables: first, how often influential moments for governments to cause trajectory changes arise, how much a funding opportunity in this cause speeds up a reform, and how likely it is that a different decision is made under the new democratic/institutional system than would have been made under the old system
- It seems like general reforms have more paths to impact, but future-focused reforms are more likely to be of enormous value across a much smaller set of paths to impact. For this reason, we think future-focused reforms are likely more important than general reforms, but are not confident in this claim (68% confidence<sup>25</sup>).
- Overall, it seems likely that there are some very important interventions in this space, but the path to impact is long and indirect and the value of many proposals seems likely to be small

## 3. Neglectedness

We might not expect democratic and institutional reforms to be a neglected cause simply because government decisions have such large and wide-ranging consequences. If the benefits

---

<sup>25</sup> This is a weird number because we took the average of our individual confidence levels (60% for SC and 75% for AG)

and costs of these decisions differ in their distribution, then many groups have strong incentives to advocate for or against changes to the political system. However, this is only true of general reforms which have effects on current outcomes. Since most beneficiaries of future-focused reforms are in the future and cannot vote or fund initiatives today, we might expect future-focused reforms to be highly neglected.

In fact, the data we've found suggest that both general and future-focused reforms seem neglected relative to their apparent importance. Future-focused reforms are particularly neglected, with practically no philanthropic funding at the moment. But while there are several large organisations advocating for general reforms, we think current philanthropic funding for these initiatives very likely totals <\$100M. That said, we haven't spent very much time looking at funding data (~2 hours total). Since democratic and institutional reform is such a large area, it's probable that we haven't found some large funders (especially outside the US).

### 3.1. Funding for general and future focused reforms

Again, we examine general and future-focused reforms separately.

#### General reforms

- Hewlett Foundation is the largest funder we've come across
  - US Democracy program awarded [\\$22M in grants in 2019](#)
  - "It supports nonprofit organizations across the ideological spectrum—academic researchers, advocacy groups, think tanks, and civic leadership organizations—who seek to understand and improve the political system so that elected representatives are better equipped to solve society's greatest problems and in turn, earn public trust and support" ([source](#))
- Wikipedia has a [list of 19 organisations in the US alone](#) working on electoral reform
  - FairVote seems to be a relatively powerful advocate for instant runoff voting, and a lowkey enemy of Center for Election Science.<sup>26</sup>
  - [Issue One](#) also seems active and well-funded
- Open Philanthropy has supported the Center for Election Science in the past, but this is not a major area of focus for them overall
  - [Total funding for CES of ~\\$2.5M](#)
  - "We see voting system reform as a neglected area with potential to facilitate more qualified candidates, increase competition and reduce hyper-partisanship in elections, and ultimately lead to improved policy decisions" ([source](#))

---

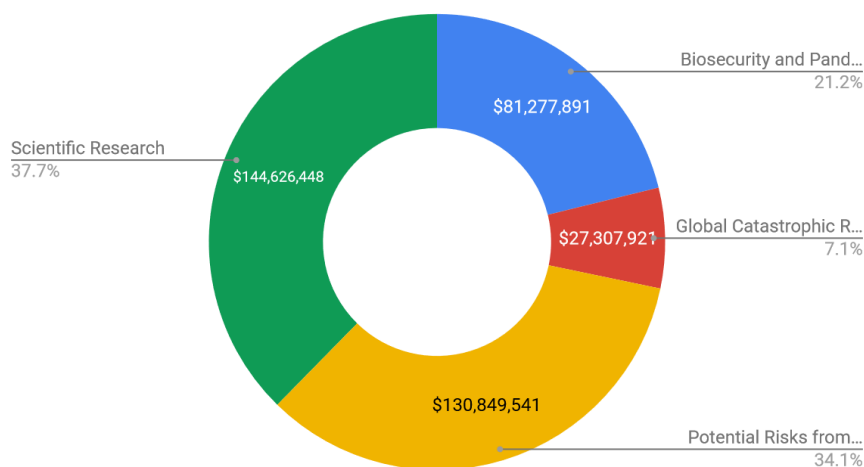
<sup>26</sup> "As more people know about instant runoff voting and are less aware of its substantial inferiority to approval voting, it has more perceived traction. Consequently, campaigns for IRV [headed by FairVote] are now following Maine's statewide implementation. As IRV is implemented, it can remove opportunity for approval voting reform within a state. It can even nullify our wins within a state. On the other side, if the instant runoff voting campaign fails or is repealed, it could sour voters to the idea of alternative voting methods altogether" ([source](#))

## Future-focused reforms

- Longview has [recommended grants](#) to found Alpenglow and support the Harvard Legal Priorities Project
  - Their grants are less open, but we expect the relevant grants total <\$5M (60% confidence)

In general, it seems that very little longtermist funding goes to future-focused electoral reforms. For example, the Open Philanthropy Project, by far the world's largest longtermist philanthropic funder, has so far overwhelmingly focused on direct existential risk reduction and field-building or accelerating scientific progress. Broad approaches aiming to influence trajectory change like democratic and institutional reform have received (practically) no funding.

Open Phil total granted up to April 2020 in longtermist causes



Open Philanthropy has funded the Center for Election Science, but our impression is that this was an unusual grant for them to make.<sup>27</sup>

## 3.2. Summary

We still have several key uncertainties about the neglectedness of this cause:

- Since research seems like a major intervention in this space, should we include the salaries of various political scientists and other researchers working on this issue?
- Should we do more work to categorize funding in terms of how relevant it is to longtermist concerns? Or divide it up using the categories we have of reforms?

Overall, our impression is that this cause does seem neglected, relative to the size of the 'problem.' The US government alone affects or directly controls how trillions of dollars are spent

<sup>27</sup> This impression is based on the grant (1) being recommended by Will Macaskill, who's not a full-time Open Phil program officer and (2) being rather unique within Open Phil's portfolio (the organisation has not made any grants to similar organisations or advocates in this space).

every year, while we're confident (70%) that <\$100M is spent per year advocating for electoral or institutional reform.<sup>28</sup> However, this seems like a good example of why we should not consider tractability and neglectedness independently.

## 4. Tractability

It's important to consider tractability in tandem with neglectedness. This is because it could simply be the case that little money is spent on an important cause because it is impossible—or at least extremely difficult—to make progress. For several reasons, we think this is likely to be the case here.

### 4.1. General considerations

In general, it seems difficult to change democratic systems because those currently in power inherently benefit from the current system, there are many different alternatives so advocates struggle to coordinate, and, in the case of future-focused reforms, it is inherently difficult to force policy makers to consider future generations because future generations cannot hold them accountable or provide feedback on policy decisions.

- **Entrenched interests benefit from status quo**
  - Politicians who have the power to champion reforms are inherently the ones who benefit from the current system, because they were elected under those rules
    - e.g. politicians elected under gerrymandered systems have strong incentives to maintain status quo; Canada's Liberal government dropped its promise to change electoral systems after it won an election under existing rules<sup>29</sup>
- **Lack of consensus on alternatives**
  - While opponents may agree that they don't like the current system, forming consensus around the best alternative is plausibly an even greater challenge<sup>30</sup>
    - e.g. Center for Election Science and FairVote both work to replace the current system in the US, but CES prefers approval voting and FairVote prefers instant runoff voting. This conflict means they compete rather than cooperate; CES's Director has said that FairVote's success comes at the expense of CES's<sup>31</sup>

---

<sup>28</sup> I'm highly confident Hewlett is the biggest funder in this space. If we assume Hewlett's \$22M accounts for just 30% of the funding in this space, then the total would be \$73M.

<sup>29</sup> <https://www.cbc.ca/news/politics/trudeau-electoral-reform-mandate-1.3961736>

<sup>30</sup> "activists for voting reform still haven't managed to use our common hatred for FPTP to unite behind a common proposal"

(<https://www.lesswrong.com/posts/D6trAzh6DApKPhbv4/a-voting-theory-primer-for-rationalists>)

<sup>31</sup> "As [instant runoff voting] is implemented, it can remove opportunities for approval voting reform within a state"

(<https://forum.effectivealtruism.org/posts/H5qhAsREAZKa3wyz7/the-center-for-election-science-appeal-for-2020>)

- Lack of consensus on alternatives was also the stated reason for electoral reform failure in Canada<sup>32</sup>
- **Inherent difficulty of incentivizing policymakers to act in the interests of future generations**
  - Proposals like Future Assemblies, Ombudsman, etc. lack strong enforcement mechanisms to hold politicians accountable to uphold their mandate

## 4.2. Empirical data on the track record of reform efforts

While empirical analyses of the success rate of reform attempts are lacking, it is our impression that such attempts fail at a high rate. We're 90% confident that the success rate is <50%, and 50% confident that it is as low as 20% or less.

- Currently 1 US state, Maine, uses instant runoff voting in its elections for governor and state and federal legislatures. It's been used since 2018.
- Alaska is supposed to start using IRV in 2022
- A law mandating the use of IRV was passed in North Carolina in 2006, but repealed in 2013
- Haven't been able to find comprehensive, public analyses of the track record of reform attempts. Pablo Stafforini analyzed attempts to get instant runoff-voting implemented at various levels in the US up to 2017. His results suggest a ~50% success rate from proposal to implementation of ballot initiatives, with ~60% still active after 6 years. [Please don't share these results externally for now.]

This is also limited to attempts to change voting systems. We haven't looked into data on efforts to change constitutions, reform bureaucratic institutions, or change laws. Our impression is that institutional reforms may be more likely to succeed, but that the significance of institutional reforms varies widely.<sup>33</sup>

## 4.3. Will it get solved by default?

It seems unlikely that this problem will get solved by default. We expect to see some general institutional reforms, but due to trade-offs between different institutional structures it is difficult to say that this problem will ever be fully "solved." Future-focused reforms seem very unlikely to be implemented without effort from longtermists, as there is no other incentive or pressure to change political systems to benefit future generations more.

---

<sup>32</sup> "the Liberal Minister of Democratic Institutions, Karina Gould, announced that a change of voting system would no longer be in her mandate, citing a lack of broad consensus among Canadians on what voting system would be best" ([https://en.wikipedia.org/wiki/Electoral\\_reform#Canada](https://en.wikipedia.org/wiki/Electoral_reform#Canada))

<sup>33</sup> For example, it seems like governments often implement institutional reforms after getting elected (e.g. creating new bodies, eliminating old bodies, combining or splitting departments, etc.), but that many of these reforms are not highly significant (especially from a longtermist point of view).

## 4.4. Summary

Our impression is that it is likely to be quite difficult to make progress on this cause. The lack of consensus on clear alternatives, resistance from entrenched interests, and difficulty of implementing enforcement mechanisms for future-focused reforms are key considerations. We can see ourselves updating this view if we saw a clearer track record of success for reform attempts.

## 5. Cost-effectiveness

We are not aware of any existing cost-effectiveness analyses of interventions or organisations in this space. Data on ballot initiatives and past attempts at initiating voting reform suggest campaigns cost on the order of \$2M to \$10M. Small-scale, future-focused advocacy can be much cheaper than this:

- Ballot initiatives cost on average ~\$2.1 million in 2020, and each signature cost ~\$8.<sup>34</sup>
- Aaron Hamlin suggests past initiatives to get IRV implemented at the state level have directly cost \$7M to \$10M per campaign<sup>35</sup>
- Looks like the APPG cost about £40,000 for a year of work (not including the costs of the time of the legislators involved)<sup>36</sup>
- Judging from the size of the organisation, I'd guess Alpenglow's budget is about £100k to £200k per year<sup>37</sup>

The relatively low costs of these initiatives suggests to me that they are likely to be cost-effective if we put any significant credence (say .1% or more) on them having a positive long-term influence. We don't think we currently know enough about this impact to judge confidently whether they are more cost-effective than other longtermist options.

## 6. What could a philanthropist do?

### 6.1. Future-focused reforms

Our impression is that the number of funding opportunities in the area of future-focused reform is currently limited. At this time it seems that a philanthropist could largely fund small, exploratory or early-stage initiatives in research and advocacy.

We think it is likely (70%) that philanthropists should focus on learning rather than directly trying to implement long lasting reforms now, for the following reasons:

---

<sup>34</sup> [https://ballotpedia.org/Ballot\\_measure\\_signature\\_costs,\\_2020](https://ballotpedia.org/Ballot_measure_signature_costs,_2020)

<sup>35</sup>

<https://forum.effectivealtruism.org/posts/H5qhAsREAZKa3wyz7/the-center-for-election-science-appeal-for-2020>

<sup>36</sup>

[https://forum.effectivealtruism.org/posts/AWKk9zjA3BXGmFdQG/appg-on-future-generations-impact-report-raising-the-profile-1#2\\_\\_Inputs](https://forum.effectivealtruism.org/posts/AWKk9zjA3BXGmFdQG/appg-on-future-generations-impact-report-raising-the-profile-1#2__Inputs)

<sup>37</sup> <https://www.alpenglow.org.uk/>

- Institutions are slow to change, so it's difficult to course-correct if reforms are misguided
- Future-focused look promising but we don't know which reforms would be best or how to implement them so that they work well (e.g. in implementing a reform, there will be very many minor decisions in the exact implementation that could significantly affect the success of the reforms)
- This suggests the value of information from further research and/or trialling proposed reforms on small scales could be high
- In the long-run, having a slightly better reform later is much better than a slightly worse one earlier (assuming the long-run is indeed long)
  - E.g. invest now with 4% annual returns vs invest in 10 years with 4.5% annual returns
  - The later investment catches up in 92 years from now
  - 300 years from now, the later investment is 4x larger
  - 3000 years from now, the later investment is almost 1 million x larger
- This all pushes towards learning over direct impact (assuming the future is long)

## Research

A philanthropist could fund research to either generate new proposals or vet existing ones. For example:

- Do research to come up with more proposals like age-weighted voting
- Analyze the strengths, weaknesses, and feasibility of proposals like age-weighted voting
- Convene researchers, e.g. at conferences, to synthesize knowledge and ideas in this area

Specific examples of promising research that could be extended include:

- [Representation of Future Generations in United Kingdom Policy-Making](#)<sup>38</sup> (supported by the [Centre for the Study of Existential Risk](#))
- [Longtermist institutional reform](#)<sup>39</sup>
- [Philosophical, institutional, and decision making frameworks for meeting obligations to future generations](#)<sup>40</sup>

## Advocacy

There are a few small groups working directly with policymakers to bring more attention to issues that affect future generations. These include:

- The [APPG](#) (supported by the [Centre for the Study of Existential Risk](#))
- [Alpenglow](#)

<sup>38</sup> Natalie Jones, Mark O'Brien, and Thomas Ryan, 'Representation of Future Generations in United Kingdom Policy-Making', *Futures*, Futures of research in catastrophic and existential risk, 102 (1 September 2018): 153–63, <https://doi.org/10.1016/j.futures.2018.01.007>.

<sup>39</sup> Tyler John and William MacAskill, 'Longtermist Institutional Reform', forthcoming, <https://philarchive.org>. See also [mindmup](#).

<sup>40</sup> Bruce E. Tonn, 'Philosophical, Institutional, and Decision Making Frameworks for Meeting Obligations to Future Generations', *Futures* 95 (1 January 2018): 44–57, <https://doi.org/10.1016/j.futures.2017.10.001>.

- [Intergenerational Foundation](#)

Advocacy-relevant research and advocacy carried out through research, such as that by [EA Sweden](#) looks potentially promising.

## 6.2. General reforms

There are many more funding opportunities in the space of general reforms. If a longtermist philanthropist presupposed that prevalent political problems like polarization imperil the potential prospects of all possible people, they could consider funding organisations working to reduce those problems:

- Organisations advocating for the implementation of (possibly) better voting systems
- Watchdog-like organisations bringing attention to contentious issues like campaign finance, gridlock, corruption, and transparent decision-making
- Organisations working to empower civil servants and policymakers to make better decisions, e.g. by providing trainings

At this time, though, the longterm impact of many of these actions seems unclear. A longtermist philanthropist could also fund research investigating the longterm implications of many of these actions. See, for example, [CSET's research](#) into whether forecasting can improve government decisions around global catastrophic risks.

## 7. Case for working on the problem

There are three main reasons to work on this cause: **it seems highly important and neglected, other longtermist funders are much less active in this space than in other longtermist causes**, and **this cause is appealing from a variety of worldviews**. We place most of the weight here on the first two considerations. The third consideration possibly provides some additional reason for Founders Pledge in particular to work on this cause.<sup>41</sup>

### 7.1. This cause seems highly important and neglected

**Future generations are neglected by existing political institutions:** Future generations aren't represented in democracies, so their well-being, rights and interests are easily neglected by current policy and decision making processes. This is partly due to political short-termism. Even actors that think beyond short-term election cycles do not explicitly think about the very long-term, so are likely scope insensitive to the size of the future. This suggests that there could be large gains for future generations by taking their well-being, rights and interests into account to a greater extent.

---

<sup>41</sup> This is because we try to influence the funding decisions of a large, diverse group of smaller funders rather than a small group of highly-aligned funders.

**Breadth:** Democratic governments make many important decisions and policies, influencing trillions of dollars per year, as well as influencing legislation that affects billions of people today, and likely many more in the future. Altering these decision and policy making processes through democratic and institutional reforms could have a large range of positive ripple effects. Democratic and institutional reforms have a lot of breadth in that there is a very wide range of potential positive ripple effects but these effects are concentrated on important levers (governmental spending, policies that affect billions or more) unlike other broad approaches (e.g. improving education). As a result, democratic and institutional reforms seem unusually promising, within the class of broad approaches to changing the future trajectory.

## 7.2. Since few other longtermist funders are focused on this cause, the best opportunities are more likely to be under-funded

**Longtermist portfolio diversification:** Longtermist worldviews can be plotted along four dimensions:<sup>42</sup>

1. Urgency of existential risk: the level of existential risk over the next century or so
2. Urgency of significant path dependencies: the extent to which the future trajectory is contingent on present/past states,<sup>43</sup> the rate of trajectory-altering decisions
3. Breadth: whether targeted or broad approaches to shaping the future are more impactful
4. Learning: how much we know and the extent to which we can take actions that foreseeably improve the future, and how this will change over time

A significant proportion of direct longtermist spending aims to mitigate specific existential risks, such as risks from AI and biorisk. This is motivated by a relatively urgent existential risk view (the next one or two centuries are very important) or a very urgent existential risk view (the next few decades are very important), in combination with a view according to which targeted attempts to mitigate specific risks are foreseeably promising. The question of how urgent other path dependencies are is relatively unimportant on such a view, as this only matters if there is a future trajectory to change (and after an existential catastrophe, there is no alterable future).

Saving or investing resources in order to have more to spend on direct work later is motivated by a patient view on both urgency dimensions and/or a view according to which our ability to foreseeably improve the future will improve over time. It does not say much about the breadth dimension.

---

<sup>42</sup> This is heavily inspired by Ben Todd's taxonomy of longtermist views. 'Ben Todd on Varieties of Longtermism and Things 80,000 Hours Might Be Getting Wrong', 80,000 Hours, accessed 5 January 2021, <https://80000hours.org/podcast/episodes/ben-todd-on-varieties-of-longtermism/>.

<sup>43</sup> Of course, existential risks are examples of extreme path dependencies. It helps to separate existential risk mitigation from other trajectory changes though.

Mitigating climate change is motivated by a relatively urgent view on existential risk and/or significant path dependencies and a relatively broad view. This is because most of the benefits from mitigating climate change flow from increased global stability and cooperation, which could be an important existential security factor and/or create path dependencies. This view seems consistent with a variety of views on the foreseeable impact dimension, but not very pessimistic views.

Democratic and institutional reforms look best on a broad view, and most likely, a view that is relatively non-urgent with respect to existential risks but according to which the future trajectory is very contingent (i.e. urgent with respect to path-dependencies, high rate of trajectory-altering decisions). It is consistent with a variety of views on the foreseeable impact dimension with a higher tolerance for pessimism than existing longtermist interventions (though not very pessimistic views). This differs quite substantially from the current longtermist portfolio of direct spending (which heavily weights urgent targeted existential risk mitigation) and the overall longtermist portfolio (which also saves/invests lots of resources). As a result, democratic and institutional reforms are promising candidates for diversifying the longtermist portfolio. How important this is depends on how much credence we assign to a broad, non-urgent with respect to existential risk and highly contingent view of longtermism.

**Longtermist EA fungeing:** Philanthropic funding for longtermist initiatives is dominated by a very small number of funders: largely Open Philanthropy alone, with some contributions from philanthropists advised by organisations like Longview Philanthropy or semi-aligned funders like Elon Musk (who funds AI safety work) and Bill Gates (who funds some biosecurity work). Relatively small actors in this space may be able to increase their impact by funding impactful initiatives in areas where they are less likely to funge with these large funders. For example, the best marginal funding opportunity in the space of democratic and institutional reform may not be as cost-effective as the marginal funding opportunity in AI safety. However, it may still be worth funding this initiative if it is more cost-effective than Open Philanthropy's expected last dollar of funding, since we are likely to funge heavily with Open Phil if we instead fund the AI safety opportunity.

If we expect Open Phil's last dollar is likely to be spent in an area where the longtermist impact is very small or zero (say, distributing malaria nets) then if funding opportunities in this space have any plausible longterm impact they could be worth funding.

### 7.3. This cause is appealing to a variety of worldviews

**Argument from democracy, equality and fairness as intrinsic values:** A case can be made for future-focused democratic and institutional reform on the basis of democracy, equality and/or fairness:

1. Democracy in a society consists in roughly equal influence over the governing of the society

2. The policies and decisions made at one time massively affect the state of the society and the way in which it is governed in later times
3. Members of the society are unable to influence the policies and decisions that are made in earlier times
4. Present members of the society have much more influence than future members over the governing of the society in future times
5. The distribution of influence over the governing of the society in future times is extremely unequal and undemocratic, significantly favouring earlier (including present) members of the society
6. Representing future generations in the governance of the society, though imperfect, would more equally share the influence of the governing of the society over time, and hence would be more democratic, equal and fair
7. The disparity in influence between the present and future generations is much larger than that within the present generation, so representing future generations presents the greatest opportunity for promoting democracy, equality and fairness

## 7.4. Summary

The case for working on this cause relies on two main points: First, future-focused reforms seem important (high potential upside) due to the large number of people potentially affected and unusually good for a broad approach as the paths to impact are concentrated on powerful levers (policymaking, governments). Second, future-focused reforms are neglected both in general and by longtermist philanthropists in particular. In combination, (1) and (2) suggest that there is potential for very high counterfactual impact.

Similarly, general reforms have some benefits relating to political short-termism and (more speculatively) broadly improving problem-solving and risk-management. Aidan believes that general reforms seem strictly less important and neglected than future-focused reforms (75% confidence), but potentially more tractable. Stephen is slightly less confident in this claim but still thinks future-focused reforms are more promising (60% confidence).

## 8. Case against working on the problem & potential risks

There are several reasons to think that this cause is not particularly promising for longtermist philanthropists. We categorize these reasons into three main categories: **doubt about the longterm impact of reforms**, **high intractability**, and **practical considerations for philanthropists**. Together, we place more weight on the first and, especially, third reasons.

## 8.1. Doubts about longterm impact

**Urgency of existential risk:** To the extent that the case for democratic and institutional reforms rests on positive trajectory changes, it also rests on existential risks being not too urgent. Otherwise, the future would be relatively short in expectation, and we would probably be better off directly mitigating existential risk. Improved democracy and institutions could be an existential security factor: reforms could be useful in mitigating existential risk, e.g. if politicians have significant influence over existential risk mitigation, or to reduce under-investment in existential risk mitigation. However, this seems likely to be less impactful than directly mitigating specific existential risks if they are urgent.

**Contingency of the future trajectory:** It could be the case that there isn't much scope for valuable trajectory changes. If, excluding existential risks, the future trajectory is not very contingent on the present (i.e. there aren't urgent, significant, non-existential-risk path dependencies on the future trajectory and trajectory-altering decisions rarely arise), then the plausibility of democratic and institutional reforms creating valuable trajectory changes is low. For example, we've hypothesised that the trajectory of the future could vary widely, depending on what our institutions and political norms are like today. This hypothesis could be misguided. There is a significant open question here about the nature of the world that reaches beyond this investigation.

**Targeted approaches vs broad approaches:** Democratic and institutional reforms are broad approaches to influencing the future but it could be the case that targeted approaches are more promising. This could be the case (1) in a world in which the unforeseeable ripple effects of reforms are not very valuable and/or they are small in number, or (2) if there is a small class of causal pathways that are most important and that should be targeted (e.g. perhaps we should focus very specifically on technical AI alignment research, biosecurity policy).

**Longtermist portfolio:** We make the case above that democratic and institutional reforms are promising candidates for diversifying the longtermist portfolio, as existing longtermist resource use is mainly guided by different longtermist worldviews to those that support democratic and institutional reform. This consideration can push the other way though: there is information contained in the fact that few longtermist resources are spent on very broad interventions trying to improve the future trajectory or mitigate existential risks. This suggests that other funders believe that other ways of influencing the long-term future are more impactful. This lack of funding could be explained to some extent by grantmaking capacity constraints rather than this kind of longtermist spending being of low impact, in which case further research and funding could be very valuable.

There is a big difference between (1) an area getting no funding because no one has had a chance to evaluate it and (2) an area getting no funding because it has been evaluated and found to be of low priority. We're inclined to think that reality is closer to (1) than (2) in this case, so this is only a weak consideration against democratic and institutional reform. For example, Open Philanthropy hasn't investigated democratic and/or institutional reforms and we haven't

found many existing analyses in our research. Additionally, broad approaches are naturally more challenging to evaluate than targeted approaches, as the effects of the former are typically harder to foresee.

**Risks of harm:** Because reforms are hard to study and their relative benefits and costs are so uncertain, a philanthropist could cause harm by advocating for a reform that turns out to be harmful once implemented. Since political systems seem hard to change, this reform could then be locked in and cause harm for years to come.

## 8.2. High intractability

**Reforming institutions is very challenging and often fails:** Attempts to reform institutions seem to fail more often than not. There are two main reasons for this. First, reforms will be resisted by entrenched interests who benefit from the status quo. Second, people who share an opposition to the current system may struggle to reach agreement on an alternative to champion. The track record of electoral reform attempts, for example, is poor because people are risk-averse and different advocates support different alternative voting systems.

**Existing anti-short-termist mechanisms:** There are already mechanisms in place to guard against political short-termism, to some extent. This suggests that some low-hanging fruit may have already been taken. It is well-understood that politicians in democracies face short-term incentives and so all advanced democracies have systems designed to mitigate this. For example, many countries have permanent civil servants that stay in their roles beyond short-term election cycles and central banks that are independent of government to guard against politically-motivated, short-termist monetary policy.

While mechanisms such as these plausibly guard against political short-termism to some extent, we expect the long-term (especially the very long-term) to be neglected by governments as such mechanisms weren't designed to improve policy and decision making on very long timescales. For example, civil servants don't think about the long-term on very long time horizons and their actions are largely determined by politicians, doing what ministers tell them to do, or more broadly following priorities set by the government. However, the extent to which civil servants' actions are tied to election-cycle bound politicians plausibly varies across policy areas, depending on how politicised they are.

## 8.3. Practical considerations for philanthropists

**There are few promising concrete funding opportunities:** At least for future-focused reforms, there are few concrete proposals, and many of the existing proposals are at an early-stage and require more research or scrutiny before they can be championed to a wider audience. This cause seems more suited to a field-building approach, which could require a longer-term commitment, than the one-off donations with constant marginal impact which better suit Founders Pledge members.

**We do not know which reforms would be best:** Relatedly, most of the proposals we've seen in this space have both benefits and downsides. There is a lack of consensus on the best approaches or initiatives to support among the sources we've read.

**General reforms:** Most of the reasons to think that future-focused reforms are important don't extend well to general reforms, as detailed in the importance section, so we expect future-focused reforms are more important (68% confidence), though with low resilience. Additionally, general reforms are more likely to be pursued and enacted without longtermist funding than future-focused reforms, as the current generation (and nearby future generations) stand to gain relatively more from general reforms.

## 8.4. Summary

The biggest concern is about whether there are concrete things to fund. The case for importance and neglectedness is fairly strong, but it's unclear what we could actually fund.

A secondary concern is about whether broad approaches are actually good. It seems plausible that existential risk targeted work is best and if so, broad approaches like this may not be that valuable.<sup>44</sup> In this case, it would be better just to focus on mitigating existential risk directly.

Another concern is whether the trajectory is sufficiently contingent (i.e. high enough background rate of trajectory-altering decisions). If not, then we're better off focusing on existential risk mitigation and/or waiting for future times with more opportunities to significantly alter the future trajectory.

## 9. Potential risks for Founders Pledge

While recognizing that work on this cause could be considered political, we do not currently think that working on this cause poses any serious risks.

## 10. What are the main uncertainties in this analysis?

- Stephen uncertainties:
  - How to think about the importance of this cause, i.e.:
    - What sort of levers on the future would reforms affect?
    - If it's stuff like ex risk through probability of war or technology regulation, how does this compare to direct work on those causes?
    - If it's trajectory changes, by how much can we shift trajectories through broad-based reforms? What are the mechanisms through which these reforms cause changes?

---

<sup>44</sup> It seems like this is Ord's view.

- If it's just generally better decision-making in the present as well as the future, how resilient are reforms to future changes? And how valuable are they on a daily basis?
  - Tangible idea of what interventions would be best and what could be funded
  - How to compare other values like democracy, justice, and fairness to more familiar (to us) values like wellbeing and lives saved
- Aidan
  - How promising are broad vs targeted approaches to improving the future?
    - I don't expect to be able to resolve this, but the case for working in this area hinges on it being at least plausible that broad approaches could be very good
  - Through what mechanisms could reforms mitigate existential risk?
    - Reforms seem more plausibly valuable through other trajectory changes but if they, in addition, could effectively mitigate existential risk, this would strengthen the case for pursuing this further
  - How do s-risks fit into this?
    - Strong democratic institutions intuitively seem pretty good for mitigating s-risks (e.g. by restricting authoritarian power, spreading political power more evenly - most people aren't evil), which could be important, but I haven't thought about this much
  - How likely is it that reforms would succeed and not be repealed?
    - We don't have much data and what we do have isn't super relevant
  - Which general reforms are most promising and what would a path to long-term impact from them look like?
    - I'm relatively sceptical about most general reforms (compared to future-focused reforms) but open to some of the best ones being quite promising
    - E.g. would general citizens' assemblies be a good step towards futures assemblies?

## 11. What should be the focus of future research?

- Benefits and path to impact of general and future-focused reforms
  - Our current analysis of the benefits and paths to impact of reforms is quite speculative
  - Because the cost-effectiveness of working on this cause involves multiplying a small chance of having impact by a large amount of impact if successful, generating a more robust estimate of the expected impact of reforms is a priority
  - This could be done by (e.g.):
    - Looking harder for existing analyses
    - Looking at comparative analyses of different voting systems/institutional arrangements in the political science literature
    - Looking at the effects of previous successful reform attempts

- Building a formal model of the path to impact that tests the effects of (1) how often opportunities for trajectory change arise and (2) how reforms affects the chance that governments make decisions that are better for future generations
- What research could a philanthropist fund
  - In the area of future-focused reforms, we think it's most likely that the best current funding opportunities for philanthropists are in research or learning
  - But we don't currently have a good sense of what specifically a philanthropist could fund here. Figuring out what the main areas of ongoing research in this area are would be a priority
- Are there opportunities to fund some of the proposed reforms on a small scale?

## 12. Overall judgements

AG: The area seems important and neglected. I'd be excited to fund promising future-focused reform research or small scale trials. My biggest concern is that it's difficult to find opportunities we could actually fund. Perhaps, focusing on existential risk mitigation would be better, but I think it's worth diversifying the longtermist portfolio if we find promising opportunities (and we're not going to settle the questions of whether to focus on existential risk mitigation vs trajectory change or broad vs targeted approaches anytime soon).

SC: I think it's unlikely this is the best cause area for Founders Pledge to focus on at this time (20% chance). The two most important considerations for me are (1) a lack of concrete funding opportunities in this space and (2) a lack of a strong argument linking specific reforms to improved long-term outcomes. On (2), I haven't seen a convincing proposal linking a specific general reform to important longtermist issues, and I have yet to see a future-focused reform that seems likely to significantly change the actions of governments. Given the field's importance and neglectedness, I think funding research or field-building in this area would be valuable, but my impression is that this requires a longer-term commitment from a funder, whereas funding from FP members is more likely to be sporadic and unpredictable. I would change my mind if I saw a more convincing case for a specific, marginal funding opportunity being likely to have a long-term impact.

## 13. Key Sources

[Longtermist Institutional Reform, John & MacAskill, forthcoming](#)

- [Comments by Jess Riedel](#)
- [Replies from John and further responses from Riedel](#)

Tobias Baumann - [Thoughts on electoral reform](#)

Tobias Baumann - [Improving our political system: An overview](#)

Tobias Baumann - [Representing future generations in the political process](#)

Caspar Oesterheld - [Summary of \*Democracy for Realists\*](#)

Sam Hilton - [Cause: Better political systems and policy making](#)