

The Utilitarian Virtues of Low-Hanging Fruit in Rich Nations – and Resulting Complications for Effective Altruism

Mark Budolfson

Version 1.81

Suppose your goal is to advise people on how best to contribute to charity - specifically, you want to advise people on how to be maximally effective utilitarian altruists, in the sense that you want to tell them how to maximize the expected amount of welfare added to the world by their donations.

With that goal in mind, one thing you might do is to investigate what advice you should give to 'normal' people, who are understood to have less than, say, \$100k to give. This is roughly what the charity evaluator GiveWell has done in recent years, and has done impressive work answering this question. With 'normal' people in mind, GiveWell's empirically-based advice tends to be to give to a small number of charities whose work tends to focus on what we might call 'boring low hanging fruit' (such as bed nets to protect people from malaria, and small cash transfers directly the world's poorest people), because for SMALL increments of additional revenue, the evidence allegedly indicates that the recommended charities in that category do more good than anything else.¹

In addition to the question of what form of charitable giving is most effective for 'normal' people, a further question is what form of charitable giving is most effective for those who can make very LARGE investments of, say, tens of millions of dollars or more. This is the big question that I'm interested in in this paper:

The Big Question. What form of charitable giving is most effective for those who can make very LARGE investments of, say, tens of millions of dollars or more.

With that introduction in mind, the outline of this paper is as follows:

In section 1, I canvass some bad answers to the Big Question given by some philanthropists.

In section 2, I articulate a better answer that is the mainstream answer from effective altruists, and explain why it seems committed to a surprisingly fine-grained expected utility estimate even though it purports to be non-quantitative.

In section 3, I note that if effective altruists were to explicitly acknowledge the points made in section 2 and adjust their discussions accordingly, this would lead to substantial gains along the dimensions of transparency and accuracy – both of which effective altruists claim to care greatly about. (Note that this is a 'minor critique', as it doesn't assume that effective altruists are really wrong about any of the advice they are giving – unlike my criticisms in what follows.)

¹ See for example the recommendations and analysis at www.givewell.org, which has emerged as the leading EA advisor. With Dean Spears, I raise some objections to GiveWell's analysis on this point in our coauthored paper, "Effective Altruism, Marginal Impact, and Fundraising: Weak Links in Effective Altruism's Chain".

In section 4, I explain some facts about a surprising possible answer to the Big Question, which is partially considered (but rejected after partial consideration) by effective altruists: soil lead abatement as a response to lead poisoning.

In section 5, I use some features of lead poisoning to explain a larger societal problem, much discussed in the recent literature, called the intergenerational transmission of inequality. I note that in developed nations, there is especially clear opportunity to address this problem with what I call ‘effective unilateral altruism’ from large donors.

In section 6, I express my dismay at rankings by effective altruists that assume, mistakenly, that progress on the intergenerational transmission of inequality is intractable. I note that changing this assumption radically changes their rankings of charities. One upshot is that there is even more reason for calculations to be transparent and subjected to stress tests by experts and the general public. Another upshot is that it appears that there are particularly exciting opportunities for effective unilateral altruism in developed nations, where the expected utility of such interventions might be higher than that of interventions elsewhere due to the very high reliability of doing good that is available domestically due to better governance capacity, mature institutions, and so on.

In section 7, I outline a competing answer to the Big Question that is suggested by the preceding discussion – namely, that LARGE donors should focus on unilateral altruism in developed nations.

In section 8, I arrive – finally – at a philosophically interesting conclusion. In this section I identify as important and surprisingly plausible, without endorsing, the following argument:

Contrarian Argument from Effective Altruism Against Aid to the Global Poor:

1. The best LARGE investments are in ‘boring low hanging fruit’ of a particular kind.
2. Boring low hanging fruit of that kind exists only in developed nations.
3. So, counterintuitively, the best LARGE investments tend to be in developed nations, even though there are far greater potential welfare gains in developing nations.

If one accepts utilitarian effective altruism as the correct fundamental aim, I believe that this argument is a surprisingly powerful one for investing one’s philanthropic giving in developed nations, especially if one is a LARGE donor. However, I also note some ways that the conclusion of this argument could be incorrect.

In section 9, I present some general objections to effective altruism’s standard use of tractability, crowdedness, and welfare potential as metrics, and I argue for use of ‘philanthropy dose response functions’ instead as a partial fix.

In section 10, I expand on some further suggestions for better metrics for effective altruism, and conclude.

1. Bad Answers to the Big Question

One obvious first answer to consider to the Big Question of what LARGE philanthropic investments have the best return from a utilitarian point of view is:

Answer 1. LARGE investments have the highest return when they are invested in the same boring low hanging fruit that has the highest return for SMALL investments – eg, bednets, cash transfers for the world's poorest, etc.

However, there are a number of reasons to think that this answer is not generally correct. First, this seems to overlook the fact that the best charities for SMALL donations don't have the capacity to scale up quickly and effectively in the way that would be required to absorb LARGE donations of tens of millions of dollars while maintaining a similar conversion ratio of welfare done per dollar given; so, giving to these charities far beyond their capacity to turn it into welfare would at best result in money sitting in their reserve accounts not doing as much good as it could be doing, or might be invested at a much lower rate of return than SMALL donations would achieve, or at worst could go to waste. Second, this answer seems to overlook the fact that single donors with deep pockets can create new charities to capture welfare gains in promising areas where no charity currently exists, and where it would thus be infeasible for normal people to coordinate in the way necessary to capture these gains – and the gains available by starting such new initiatives are sometimes much larger than the gains available by investing the same amount in any existing charity. Third, a single individual with deep pockets can sometimes leverage more gains via strategic moves than can existing traditional charities with a similar bankroll that are funded by normal people.

With these considerations in mind, a common alternative answer is:

Answer 2. LARGE investments have the highest return when they are invested in high risk / high reward initiatives in areas where the most welfare is at stake.

This is a principle that sometimes seems to appeal to leaders of foundations, universities, and so on insofar as those leaders' investments are made with the goal of maximally benefiting society. However, a problem with this idea as formulated and as actually practiced is that it is far too promiscuous in its endorsement of ill-conceived initiatives with basically no prospect of success that happen to be in domains with large potential welfare gains.

For example, climate change is currently everyone's favorite example of an area with large potential welfare gains, and as a result of this it follows almost by definition that *any speculative investment in a climate change initiative of any sort counts as a 'high risk / high reward investment'*. As a result of this together with endorsement of the investment principle under discussion, it is easy to find, at any major research university, examples of multi-million dollar climate change initiatives funded by well-intended donor money that have no likelihood of making any positive difference or adding any useful knowledge to the world at all. The point I want to make here with this example of climate change is NOT AT ALL that there is anything undesirable about massive funding of promising work on climate change (after all I'm in FAVOR of that); rather, my point is that because any speculative initiative on climate change whatsoever will count as a high risk / high reward initiative, too many bad initiatives that have no prospect of benefiting society at all are funded simply because they automatically count as 'high risk / high reward investments' to leaders at many universities and foundations, who seem to be operating with the sort of overly crude and overly promiscuous investment principle in mind that Answer 2 represents.

I think another factor here is what we might call an investment bias toward 'the most exciting things that will be the most important going forward in the judgment of the effective altruism community',

which often leads to overinvestment in those things. Possible examples from effective altruists include what I take to be overinvestment in synthetic meat, existential risks, and the like – ie, the faddish ‘most exciting things’ that all advocates of effective altruism talk about all the time.² An example from Warren Buffett illustrates how similar biases are in play in investment decisions more generally: according to Buffett, collectively, the airline industry has never made any money in its entire history. But for much of the 20th century, it was nevertheless the cool, high-risk, high-reward investment, and ‘one of the most important things going forward’, so it attracted a lot of overinvestment because of groupthink and bias. I would say the same thing is true of much current investment in synthetic meat, existential risks, and the like. Another reason to be skeptical of those initiatives is that I don’t think I’ve ever heard their proponents within EA identify any clear interventions that could be taken that seem like plausible ways of reducing existential risks. But I don’t want to dwell on this here. That’s an important topic for an entire EA paper of its own, but is not my main concern here. These comments merely indicate why I do not focus more on those topics here or elsewhere.

2. The Standard Answer Given by Effective Altruists, and How it Implicitly Purports to be an Estimate of Expected Utility

With the problems with Answers 1 and 2 in mind, an alternative that seems quite promising is this answer to the Big Question of what LARGE investments are best:

Answer 3. LARGE investments have the highest return when they are invested in new initiatives in domains where there are (a) large potential welfare gains, (b) it is realistic that those gains can be captured, and (c) few others are pursuing those gains – in short, areas where there’s high **welfare potential**, high **tractability**, and low **crowdedness**.³

As an example of philanthropy guided by this kind of principle, we can examine a recent summary chart from Good Ventures that summarizes its ranking of opportunities for LARGE investments in domestic initiatives using this methodology.⁴ Good Venture’s overall rankings are listed on the left, where Priority = 1 is the highest ranking:

² For some discussion, see Dylan Matthews, “I spent a weekend at Google talking with nerds about charity. I came away...worried”, online at <http://www.vox.com/2015/8/10/9124145/effective-altruism-global-ai>.

³ See the rankings of leading EA evaluators such as available at goodventures.org, givingwhatwecan.org, and Will MacAskill. 2015. *Doing Good Better*. It is striking how homogenous the top effective altruism charities are in their agreement with this answer to the big question.

⁴ Again, I’m going to ignore their evaluation of non-domestic policy initiatives for reasons that will become clear much later, and because they tend to be focused on specific things like existential risk that, as indicated above, I think are already overinvested in by effective altruists.

	Priority	Cause	Welfare Potential		Tractability		Crowdedness	
	1	Criminal justice reform	Moderate		Unusual window		Moderate	
	1	Macroeconomic stabilization	High		Highly uncertain		Low	
	2	Labor mobility	Very high		Appears low		Very low	
	2	Land use reform	Moderate		Highly uncertain		Very low	
	2	Factory farming	Depends heavily on value j		Highly uncertain		Low	
	2	Marijuana policy	Low-moderate		Unusual window		Low	
	3	Foreign aid policy	High		Highly uncertain		Moderate	
	3	Soil lead reduction	Moderate		Highly uncertain		Very low	
	4	Tax policy	High		Appears low		Moderate	
	4	Alcohol policy	Moderate		Highly uncertain		Low	
	4	Organ donation	Low-moderate		Highly uncertain		Very low	
	5	State and local advocacy/pol	High		Highly uncertain		Uncertain	
	5	Paid sick days	Moderate		Highly uncertain		Moderate	
	5	Wage theft	Moderate		Highly uncertain		Moderate	
	6	Intellectual property reform	Moderate		Highly uncertain		Moderate	
	6	Other redistribution/inequalit	High		Highly uncertain		Moderate	
	6	Health care policy	High		Highly uncertain		High	
	7	Low-skill occupational licens	Low-moderate, maybe low		Highly uncertain		Low	
	7	Civil legal assistance	Low-moderate, maybe low		Highly uncertain		Moderate	
	7	Other rent seeking	Highly uncertain		Highly uncertain		Uncertain	
	7	Improving democracy	High		Highly uncertain		High	
	8	Broadband policy	Low-moderate		Highly uncertain		Moderate	
	8	Climate change	High		Highly uncertain		High	

Again, everything on this chart is taken from the Good Ventures site. The first thing I want to note is that the overall rankings in the left-most ‘Priority’ column can be seen as a more or less a function of rankings along the dimensions of welfare potential, tractability, and crowdedness that are implicit in the seemingly qualitative ratings in those columns.

To bring this out, in the next graph I’ve introduced a function that converts in a straightforward way the qualitative rating along each of those dimensions to a numerical value between 1 and 10 that merely presents an ordinal ranking of how good the qualitative rating indicates that the cause does along that dimension, where 10 is doing best. Then in the new leftmost ‘Product’ column, I multiply the numerical ratings along each of these dimensions together to get an aggregate rating, where a higher score is better.

Product	Priority	Cause	Welfare Potential	W	Tractability	T	Crowdedness	C
120	1	Criminal justice reform	Moderate	5	Unusual window of	8	Moderate	3
112	1	Macroeconomic stabilization	High	8	Highly uncertain	2	Low	7
100	2	Labor mobility	Very high	10	Appears low	1	Very low	10
100	2	Land use reform	Moderate	5	Highly uncertain	2	Very low	10
70	2	Factory farming	Depends heavily on	5	Highly uncertain	2	Low	7
112	2	Marijuana policy	Low-moderate	2	Unusual window of	8	Low	7
48	3	Foreign aid policy	High	8	Highly uncertain	2	Moderate	3
100	3	Soil lead reduction	Moderate	5	Highly uncertain	2	Very low	10
24	4	Tax policy	High	8	Appears low	1	Moderate	3
70	4	Alcohol policy	Moderate	5	Highly uncertain	2	Low	7
40	4	Organ donation	Low-moderate	2	Highly uncertain	2	Very low	10
48	5	State and local advocacy/pol	High	8	Highly uncertain	2	Uncertain	3
30	5	Paid sick days	Moderate	5	Highly uncertain	2	Moderate	3
30	5	Wage theft	Moderate	5	Highly uncertain	2	Moderate	3
30	6	Intellectual property reform	Moderate	5	Highly uncertain	2	Moderate	3
48	6	Other redistribution/inequalit	High	8	Highly uncertain	2	Moderate	3
16	6	Health care policy	High	8	Highly uncertain	2	High	1
28	7	Low-skill occupational licens	Low-moderate, mayt	2	Highly uncertain	2	Low	7
12	7	Civil legal assistance	Low-moderate, mayt	2	Highly uncertain	2	Moderate	3
12	7	Other rent seeking	Highly uncertain	2	Highly uncertain	2	Uncertain	3
16	7	Improving democracy	High	8	Highly uncertain	2	High	1
12	8	Broadband policy	Low-moderate	2	Highly uncertain	2	Moderate	3
16	8	Climate change	High	8	Highly uncertain	2	High	1

As you can see, this seems to capture pretty well how Good Ventures takes the overall ranking of these causes to be a function of its rating along each of these dimensions. Note that nothing I've done here requires us to think that the numbers here 'mean anything' in the sense of 'quantifying' any of the underlying relevant facts beyond the ordinal rankings that are obviously implicit in them. On the contrary, so far the numbers here are merely a restatement of the orderings that are already implicit in the qualitative claims that some causes do better and worse than others in an orderable way along the dimensions of welfare potential, tractability, and crowdedness. So, reasoning about these numbers does not presuppose any problematic ability to quantify anything beyond what is already implicit in Good Ventures's qualitative ratings. And I think you'll find, as I did, that changing these numbers up slightly, or the way of aggregating them, does not make much difference to the aggregate results you see here, so long as the numbers you choose are more or less consistent with the ordering that is implicit in the qualitative ratings along each dimension.⁵

The next thing I want to do is multiply the numbers in the welfare potential column by 100, and multiply the numbers in the tractability and crowdedness columns each by 1/10; if we do that, then multiplying the columns together still gives us the same product:

⁵ The spreadsheet containing all of the calculations throughout this paper (and the original data from Good Ventures) is available at [www.budolfson.com/Budolfson Good Ventures.xlsx](http://www.budolfson.com/Budolfson%20Good%20Ventures.xlsx), based on a spreadsheet downloaded in 2014 from Good Ventures at <http://www.openphilanthropy.org/focus/us-policy>.

Product	Priority	Cause	Welfare Potential	W	Tractability	T	Crowdedness	C
120	1	Criminal justice reform	Moderate	500	Unusual window of	0.8	Moderate	0.3
112	1	Macroeconomic stabilization	High	800	Highly uncertain	0.2	Low	0.7
100	2	Labor mobility	Very high	1000	Appears low	0.1	Very low	1
100	2	Land use reform	Moderate	500	Highly uncertain	0.2	Very low	1
70	2	Factory farming	Depends heavily on	500	Highly uncertain	0.2	Low	0.7
112	2	Marijuana policy	Low-moderate	200	Unusual window of	0.8	Low	0.7
48	3	Foreign aid policy	High	800	Highly uncertain	0.2	Moderate	0.3
100	3	Soil lead reduction	Moderate	500	Highly uncertain	0.2	Very low	1
24	4	Tax policy	High	800	Appears low	0.1	Moderate	0.3
70	4	Alcohol policy	Moderate	500	Highly uncertain	0.2	Low	0.7
40	4	Organ donation	Low-moderate	200	Highly uncertain	0.2	Very low	1
48	5	State and local advocacy/pov	High	800	Highly uncertain	0.2	Uncertain	0.3
30	5	Paid sick days	Moderate	500	Highly uncertain	0.2	Moderate	0.3
30	5	Wage theft	Moderate	500	Highly uncertain	0.2	Moderate	0.3
30	6	Intellectual property reform	Moderate	500	Highly uncertain	0.2	Moderate	0.3
48	6	Other redistribution/inequalit	High	800	Highly uncertain	0.2	Moderate	0.3
16	6	Health care policy	High	800	Highly uncertain	0.2	High	0.1
28	7	Low-skill occupational licens	Low-moderate, mayt	200	Highly uncertain	0.2	Low	0.7
12	7	Civil legal assistance	Low-moderate, mayt	200	Highly uncertain	0.2	Moderate	0.3
12	7	Other rent seeking	Highly uncertain	200	Highly uncertain	0.2	Uncertain	0.3
16	7	Improving democracy	High	800	Highly uncertain	0.2	High	0.1
12	8	Broadband policy	Low-moderate	200	Highly uncertain	0.2	Moderate	0.3
16	8	Climate change	High	800	Highly uncertain	0.2	High	0.1

At this point, all that has really happened is that I've done a simple transformation on the original numbers that could be interpreted as merely a non-standard representation of the same initial idea that along each dimension there is a quantifiable ranking implicit in the qualitative ratings that Good Ventures reports – and again, multiplying the new values here together still gives us the same product. With all of this in mind, the transformation I've performed here is consistent with maintaining that these numbers still do not really mean anything.

However, at the same time, there is now an obvious interpretation of these numbers on which they do mean something, and on which they mean something very specific that is of crucial importance to the goal of properly ranking charities from a utilitarian point of view. That's because it is now easy to give a gloss on what is meant by these numbers, which I think is also implicit in Good Ventures's ratings, that adds up to a judgment about the expected utility associated with each of these investment options – which is, after all, the ultimate criterion of better and worse LARGE investments that Good Ventures wants its ratings to track:

E(Utility)	Priority	Cause	Welfare Potential	W	Tractability	T	Crowdedness	C
120	1	Criminal justice reform	Moderate	500	Unusual window of	80%	Moderate	30%
112	1	Macroeconomic stabilization	High	800	Highly uncertain	20%	Low	70%
100	2	Labor mobility	Very high	1000	Appears low	10%	Very low	100%
100	2	Land use reform	Moderate	500	Highly uncertain	20%	Very low	100%
70	2	Factory farming	Depends heavily on	500	Highly uncertain	20%	Low	70%
112	2	Marijuana policy	Low-moderate	200	Unusual window of	80%	Low	70%
48	3	Foreign aid policy	High	800	Highly uncertain	20%	Moderate	30%
100	3	Soil lead reduction	Moderate	500	Highly uncertain	20%	Very low	100%
24	4	Tax policy	High	800	Appears low	10%	Moderate	30%
70	4	Alcohol policy	Moderate	500	Highly uncertain	20%	Low	70%
40	4	Organ donation	Low-moderate	200	Highly uncertain	20%	Very low	100%
48	5	State and local advocacy/pov	High	800	Highly uncertain	20%	Uncertain	30%
30	5	Paid sick days	Moderate	500	Highly uncertain	20%	Moderate	30%
30	5	Wage theft	Moderate	500	Highly uncertain	20%	Moderate	30%
30	6	Intellectual property reform	Moderate	500	Highly uncertain	20%	Moderate	30%
48	6	Other redistribution/inequalit	High	800	Highly uncertain	20%	Moderate	30%
16	6	Health care policy	High	800	Highly uncertain	20%	High	10%
28	7	Low-skill occupational licens	Low-moderate, mayt	200	Highly uncertain	20%	Low	70%
12	7	Civil legal assistance	Low-moderate, mayt	200	Highly uncertain	20%	Moderate	30%
12	7	Other rent seeking	Highly uncertain	200	Highly uncertain	20%	Uncertain	30%
16	7	Improving democracy	High	800	Highly uncertain	20%	High	10%
12	8	Broadband policy	Low-moderate	200	Highly uncertain	20%	Moderate	30%
16	8	Climate change	High	800	Highly uncertain	20%	High	10%

The basic idea here is that the Welfare Potential column tells us the amount of welfare that *could* be captured in the domain referred to by the ‘Cause’ column; then, the ‘Tractability’ column tells us how likely it is that these potential welfare gains *can actually be captured* (given constraints of political feasibility and the like); then, the ‘Crowdedness’ column tells us, conditional on these gains being captured in part by our LARGE investment, how many other philanthropy players will have been involved in capturing them, and thus *what percentage of the captured gains will be attributable to our LARGE investment*. (Here I’ve converted the decimal values for tractability and crowdedness from the last graph to percentage values.)

If we endorse the idea that each of these columns has this meaning, then we should also agree that the result of multiplying them together means something too, and in particular, that it represents an estimate of the expected utility of a LARGE investment of a particular size in each of these domains.⁶ Again, the idea is that we take the amount of potential welfare gain, multiply it by the chance that that potential can actually be captured, and then multiply that by the percentage of the capture that would be attributable to our LARGE investment if it were actually captured; the result is, very roughly, an estimate of the expected welfare gain from a LARGE investment in that domain.

⁶ The idea of what percentage of the gains are attributable to a particular investment raises further problems even at the level of these broad utilitarian brush strokes, as this idea does not track the notion of marginal effect that should be the focus of correct utilitarian analysis, but I will ignore this issue here because this kind of mistake is implicit in almost all EA thinking about these issues. I raise some objections along these lines in the last two section below. For more discussion of this, see “Effective Altruism, Marginal Impact, and Fundraising: Weak Links in Effective Altruism’s Chain”, coauthored with Dean Spears.

3. The Inescapability of Serious Quantitative Rankings, a Transparency Deficit, and Minor Suggestions

One thing this shows is that if you're Good Ventures, you cannot easily get out of a commitment to quantitative claims simply by refusing to make them, unless you also give up on a number of other things that I think Good Ventures should not give up on. In particular, I think there are two essential things that they should not give up on that are sufficient to commit them to some quantitative claims about expected utility. First, I think they'd be unwilling to give up on making comparative ratings along each of the dimensions that they think are most important in a way that amounts to a cardinal ranking along each of those dimensions. (Cardinal in the sense that, for example, there is a much larger difference between 'low moderate' and 'very high' welfare potential than there is between 'moderate' and 'high'.) Second, I think that they should be unwilling to give up on having their all things-considered evaluation be a principled function of their ratings along these dimensions. But as we've seen, if the only dimensions of evaluation at issue are essentially estimates of welfare and the probabilities relevant to capturing that welfare (as makes sense from a utilitarian point of view), then this seems sufficient to generate some substantive quantitative claims about expected utility, such as for example the claim that "the expected utility of a LARGE investment in criminal justice reform is several times greater than a similar investment in intellectual property reform".⁷

I don't see this as a problem with their approach; in fact, I see it as a virtue that what they are doing can be interpreted as estimating expected utility. But my inclination is to think that we should take things here a step further and make what is implicit more explicit. More specifically, I think that once we see that we are unwilling to give up on a scheme of evaluation within which such rankings are implicit both at the level of each relevant dimension and at the aggregate level, it then seems that such a scheme of evaluation will only be improved by making such rankings explicit. At the very least, this would improve things along the dimension of *transparency*, which they claim to hold as one of their primary values.

Note that GiveWell, Open Philanthropy, and Good Ventures (which are staffed and managed by essentially the same people) claim to have transparency as one of their very top objectives; for example, here is from the GiveWell website: "We maintain pages explaining the full details of our process for identifying top charities, along with lists of all charities considered which links to our explanation for why these charities do not qualify for our top ratings (subject to charities asking us to keep materials confidential)." Note that they are not living up to this in connection with the analysis I'm discussing, as there seem to be further calculations that they take to justify the stuff on these sheets that they choose not to display on their website. As one illustration, note that a ranking based on the numbers I've added makes more sense to be than Good Venture's own rankings on the spreadsheet – for example, 'Foreign Aid Policy' and 'Other Redistribution/Inequality' have the same ratings along every dimension, but the former is given a much higher ranking than the latter; similar remarks apply to 'Land Use Reform' and 'Soil Lead Reduction', and so on. So, it appears that there must be some highly opaque further

⁷ In other words, it then generates some cardinal claims about the comparative goodness and badness of options, with of course some degree of imprecision (which derives from the imprecision involved in the question of what exact cardinal rankings are implicit in the qualitative evaluations of each dimension). The claim is not that it commits one to any view about the particular magnitude of expected good that would be done by any option.

considerations applied to make it the case that those pairs of causes do not get the same priority rankings.

I suspect that using an explicit estimate of expected utility would also improve things along other dimensions such as *accuracy*, or at least not do harm along those dimensions. For example, if the qualitative ratings were replaced by both explicit estimates of welfare and explicit estimates of the relevant probabilities, then the spreadsheet could be more readily reviewed and critiqued by economists and other experts, who could then more readily identify research and other considerations that support or tell against the evaluation of interventions in the relevant domains. Similarly, they could then more easily present reasons for thinking that other domains not yet considered by Good Ventures had more promising expected welfare profiles than those currently under consideration.

4. A Surprise Contender: Soil Lead Abatement

Let's now turn our focus to the highest rated causes that Good Ventures ranks, and then focus on the one that may initially appear to be out of place, which I think provides a useful example.

E(Utility)	Priority	Cause	Welfare Potential	W	Tractability	T	Crowdedness	C
120	1	Criminal justice reform	Moderate	500	Unusual window of	80%	Moderate	30%
112	1	Macroeconomic stabilization	High	800	Highly uncertain	20%	Low	70%
100	2	Labor mobility	Very high	1000	Appears low	10%	Very low	100%
100	2	Land use reform	Moderate	500	Highly uncertain	20%	Very low	100%
70	2	Factory farming	Depends heavily on	500	Highly uncertain	20%	Low	70%
112	2	Marijuana policy	Low-moderate	200	Unusual window of	80%	Low	70%
100	3	Soil lead reduction	Moderate	500	Highly uncertain	20%	Very low	100%

If you're like most people, when you think about the most promising opportunities for improving society, soil lead reduction does not come to mind even as a contender, unlike the other causes on this list, which at the very least promise some obviously large upside if we move things in a progressive direction. I myself was surprised when I looked at the Good Ventures website and noticed they had chosen lead abatement as worthy of close attention, let alone as a ranked intervention.

However, this came as a *pleasant* surprise to me, because I think it shows great openmindedness and insight on the part of Good Ventures, as well as resistance to the biased conception of 'the most exciting things that will be the most important going forward' in the minds of effective altruists, as discussed earlier.⁸ This is true for a number of reasons that are not widely known, and to some extent have only come clearly into focus recently from excellent empirical research by economists and other experts. I'm going to quickly summarize the evidence here as I understand it, because it is a good illustration of some claims I want to rest weight on later in this paper, where I will go on to give some arguments that are of more general importance to effective altruism.

So, one reason lead in soils is important may be somewhat familiar: lead causes cognitive problems, which are largely irreversible when lead poisoning causes developmental problems in children. What is

⁸ The fact that Good Ventures is thinking carefully about this when no one else is just one example of the great virtues of their work.

not generally recognized is that lead *in soil* is arguably a bigger problem than both lead paint and lead in water combined.⁹

Why is there lead in soil? Part of the story is lead from old lead paint chips. But a bigger part of the story is that there used to be lead in gasoline, which spewed lead into the air for decades, which then settled into the adjacent soil – and this lead accumulation in soil was of course especially concentrated along roadways that were heavily used, which are places that even today are disproportionately inhabited by the poor.¹⁰ And so now, at various times of the year, lead in soil becomes more easily airborne, where it can then be ingested by breathing it in; alternatively, children who play outside in many urban areas can obtain a dangerous dose of lead simply from licking their hands once a day; as another alternative, this lead can be tracked into houses on the bottom of one's shoes, where it can then become airborne or ingested in some other way; finally, some vegetables grown in inner city gardens and other contaminated areas constitute a highly efficient means of transferring lead from the soil into human bodies, or more generally through soil that is not completely washed off of these vegetables.

5. The Intergenerational Transmission of Inequality, and Opportunities for Effective Unilateral Altruism in Developed Nations

Lead poisoning is merely one piece of a more general and enormously important social problem that has recently been given the name of 'the intergenerational transmission of inequality'.¹¹ The current thinking about this is that many desirable cognitive abilities and behavioral dispositions require proper development of our brains very early in our lives to achieve – in many cases, they require proper development in the womb before we are even born. Furthermore, this development is importantly undermined if very early in our lives we have suboptimal nutrition, exposure to toxins, or exposure to stress hormones in the womb. Sadly, poor people, in virtue of being poor, are far more likely to be exposed to all of these detrimental factors that prevent proper development, and if they are exposed to these things very early in life, this permanently handicaps them throughout their lives relative to the average more affluent member of society. The nature of these handicaps is such that they can never really be removed later no matter how much money we pour into education and other social welfare initiatives.

As a result of all of this, there is a causal connection between being conceived by parents at the bottom of the socioeconomic ladder and ending up there later in life, even when other factors are held fixed. And the fact that these handicaps are fixed for all time by events very early in one's life implies that this is a particularly thorny problem that cannot be solved simply by slightly larger redistributions of wealth of the kind currently under discussion.

⁹ Mielke and Reagan. 1998. "Soil is an important pathway of human lead exposure", *Environmental Health Perspectives*.

¹⁰ Substantial contributions to lead in soil have also been made by exterior lead based paint, and other industrial emitters. Lead contamination from all of these sources is additive.

¹¹ See Anna Aizer and Janet Currie. 2014. "The intergenerational transmission of inequality", *Science*, and the references therein.

Fortunately, there is one small bit of good news in all of this, which is that unlike many other big social problems, in developed nations with high governance capacity, philanthropy can make significant progress on many of the causes of the intergenerational transmission of inequality through *unilateral* action that has a very high probability of achieving the desired welfare gains, partly because in order to succeed it doesn't require the cooperation of legislators or entrenched special interest groups that would be required for other similarly large welfare gains on other social problems. For example, if there is lead in soils in an inner city area, it is feasible simply to pay the owners of this often derelict property to allow a foot of clean topsoil and a layer of sod to be put on top of the existing soil, which has the effect of largely sequestering the lead.¹² This is something that philanthropists can simply do *unilaterally*, and it is highly *scalable*, in the sense that there are a lot of large cities in the United States with a lot of acreage that is ripe for lead abatement. Furthermore, given that some municipalities have already taken similar remediation action themselves, it appears highly feasible to enlist their support in incentivizing cooperation by property owners with further remediation if a charity is willing to pick up the tab for this incentivizing. Finally, soil lead abatement also has other important 'co-benefits', in the sense that other pollutants in the existing top soil would also be sequestered as a welcome side effect of sequestering the lead.

Somewhat analogous remarks apply to many other toxins that play a role in the intergenerational transmission of inequality, especially those delivered through air pollution, where there is often existing regulatory authority (in particular, under the EPA's clean air act authority) that can be used to reduce these pollutants without the need to pass any new legislation. And economic analyses indicate that for many air pollutants there would be very large welfare gains from such reductions. For example, a recent economic analysis of the costs and benefits of an optimal reduction of the most familiar air pollutants (namely, SO₂ and particulate emissions) estimates over a trillion dollars in domestic U.S. gains from modifications to existing regulation that would ratchet down and otherwise optimize the existing regulations of these pollutants, and which would, again, not require any new legislation to authorize.¹³ Of course, unlike lead in soils, further mitigation of these air pollutants would require a large investment to overcome special interest opposition, but because the executive branch controls the EPA and because new regulation would not require new authority or legislation, it appears much more feasible to get new regulations enacted here than in other areas where similarly large welfare gains would require new redistributive legislation.

With the large social problem of the intergenerational transmission of inequality and its causes in mind, it is excellent that Good Ventures has on its radar initiatives like soil lead reduction. Beyond soil lead reduction there are a few related initiatives that have somewhat similar virtues of scalability and I would argue a clear advantage over most other initiatives under discussion along the dimension of tractability, where again, this is true because these highlighted initiatives can be achieved more or less through unilateral action that doesn't require additional legislation, unlike the others. I list a few others here for the sake of discussion:

¹² When lead levels are especially high, the top foot of existing soil can also be removed before fresh soil is brought in.

¹³ For example, Muller and Mendelsohn. 2012. *Using Marginal Damages in Environmental Policy: A Study of Air Pollution in the United States*, and Muller, Mendelsohn, and Nordhaus. 2011. "Environmental Accounting: Methods with an Application to the United States Economy", *American Economic Review*.

E(Utility)	Priority	Cause	Welfare Potential	W	Tractability	T	Crowdedness	C
120	1	Criminal justice reform	Moderate	500	Unusual window of	80%	Moderate	30%
112	1	Macroeconomic stabilization	High	800	Highly uncertain	20%	Low	70%
100	2	Labor mobility	Very high	1000	Appears low	10%	Very low	100%
100	2	Land use reform	Moderate	500	Highly uncertain	20%	Very low	100%
70	2	Factory farming	Depends heavily on	500	Highly uncertain	20%	Low	70%
112	2	Marijuana policy	Low-moderate	200	Unusual window of	80%	Low	70%
100	3	Soil lead reduction	Moderate	500	Highly uncertain	20%	Very low	100%
		EPA Lower Pollution						
		WIC Supplement						

6. The Fragility of GoodVenture's Rankings, and a Place for Stress Testing

In light of the points made in the previous section, I have to say I'm a little bit puzzled by Good Venture's rating of soil lead reduction as having 'highly uncertain' tractability:

E(Utility)	Priority	Cause	Welfare Potential	W	Tractability	T	Crowdedness	C
120	1	Criminal justice reform	Moderate	500	Unusual window of	80%	Moderate	30%
112	1	Macroeconomic stabilization	High	800	Highly uncertain	20%	Low	70%
100	2	Labor mobility	Very high	1000	Appears low	10%	Very low	100%
100	2	Land use reform	Moderate	500	Highly uncertain	20%	Very low	100%
70	2	Factory farming	Depends heavily on	500	Highly uncertain	20%	Low	70%
112	2	Marijuana policy	Low-moderate	200	Unusual window of	80%	Low	70%
100	3	Soil Lead reduction	Moderate	500	Highly uncertain	20%	Very low	100%
		EPA Lower Pollution						
		WIC Supplement						

I mean, is the idea here that reforming the criminal justice system is a piece of cake compared to the task of paying people to dump some dirt and lay some sod? (Note that this seems to be what is implied by their ratings here.)

In this particular case, it is hard for me to believe that Good Ventures' assessment of this is correct. At the very least, let's see what happens when we leave all of Good Venture's ratings exactly as they are but merely change their rating for the tractability of soil lead removal so that it is at least as good as their rating for the tractability of criminal justice reform:

E(Utility)	Priority	Cause	Welfare Potential	W	Tractability	T	Crowdedness	C
120	1	Criminal justice reform	Moderate	500	Unusual window of	80%	Moderate	30%
112	1	Macroeconomic stabilization	High	800	Highly uncertain	20%	Low	70%
100	2	Labor mobility	Very high	1000	Appears low	10%	Very low	100%
100	2	Land use reform	Moderate	500	Highly uncertain	20%	Very low	100%
70	2	Factory farming	Depends heavily on	500	Highly uncertain	20%	Low	70%
112	2	Marijuana policy	Low-moderate	200	Unusual window of	80%	Low	70%
400	3	Soil Lead reduction	Moderate	500	Good	80%	Very low	100%
		EPA Lower Pollution						
		WIC Supplement						

As you can see, merely making this change – which I see as merely making a correction – has the implication that LARGE investments in soil lead reduction are now estimated to have expected utility several times higher than investments in any of the other top-ranked initiatives.

If we add fairly conservative estimates for reductions in air pollution through improved EPA regulation as well as the sort of conditional cash transfer add-on to the WIC program discussed above, it looks like lowering air pollution through the EPA should also score well:

120	1	Criminal justice reform	Moderate	500	Unusual window of	80%	Moderate	30%
112	1	Macroeconomic stabilization	High	800	Highly uncertain	20%	Low	70%
100	2	Labor mobility	Very high	1000	Appears low	10%	Very low	100%
100	2	Land use reform	Moderate	500	Highly uncertain	20%	Very low	100%
70	2	Factory farming	Depends heavily on	500	Highly uncertain	20%	Low	70%
112	2	Marijuana policy	Low-moderate	200	Unusual window of	80%	Low	70%
400	3	Soil Lead reduction	Moderate	500	Good	80%	Very low	100%
250		EPA Lower Pollution	Moderate	500	Decent	50%	Very low	100%
48		WIC Supplement	Low-moderate	200	Good	80%	Moderate	30%

As further relevant commentary, I should mention that I haven't even yet discussed one of the most important benefits of reducing lead in soils, which is related to reductions in crime. To get a feel for the magnitude of the potential welfare gain here, some economists have recently argued that lead reductions are responsible for over 50% of the reductions in violent crime both in the USA and other countries in the last several decades.¹⁴ There is a clear, well-understood mechanism that makes sense of how this could be true, which is that even fairly low levels of lead exposure are known to increase aggressiveness and measurably decrease capacities to control impulses.¹⁵ More surprising though is the large *magnitude* of crime reduction that lead reduction seems to explain, and as a consequence the amount of further welfare gains through further reductions in violent crime that could be achieved by further reductions in soil lead levels. To get a feel for the regression analyses that are the basis for the claim that over 50% of crime reductions are attributable to reductions in lead levels, imagine that you look carefully at the exact times that the phase out of leaded gasoline happened at different locations across the US and around the world (which is an interesting thing to look at because leaded gasoline was phased out at different times at different locations, even within the US, where the most precise data is available). When you look at this, you discover that the timing and magnitude of reductions in aggregate lead emissions are followed with a very predictable short time lag by reductions in violent crime, the magnitude of which is also predicted by the magnitude of the lead reductions – and more specifically, by econometric analysis, you find that over 50% of total crime reductions are explained by lead reductions and cannot be explained by other plausible factors like increased policing, increased access to abortion, and other factors commonly hypothesized to explain the reduction in violent crime.¹⁶

¹⁴ Jessica Wolpaw Reyes, "ENVIRONMENTAL POLICY AS SOCIAL POLICY? THE IMPACT OF CHILDHOOD LEAD EXPOSURE ON CRIME", NBER

¹⁵ See Reyes pg 48 for references to numerous studies over many decades.

¹⁶ See Reyes, throughout.

With this in mind, it is interesting to note that the potential welfare gains from soil lead reduction might be much higher than they appear even after an initial foray into the issue of lead in soil. And in ongoing work, Janet Currie and collaborators are, using excellent methodologies, finding substantial connections between lead reductions and improved educational outcomes, further solidifying the connection between lead exposure and foregone wellbeing, and the perpetuation of deep social problems via the intergenerational transmission of inequality.¹⁷ I think similar remarks also apply in the case of air pollution, although I lack the space to go into the details here.¹⁸

As a full disclosure, I should say that one of my research interests is air pollution control and, more generally, environmental quality and regulation, which means that I'm undoubtedly biased in favor of thinking that environmental issues are more important than they actually are relative to other things. But having said that, I think the causal factors brought together under the narrative of the intergenerational transmission of inequality are very real, are underappreciated, and in many cases represent very tractable and underexploited opportunities for effective altruism – and that's all I need for the basic points I want to make here, together with the fact that reliable progress on these fronts is only available in developing nations via unilateral philanthropy, due to institutional and governance facts. I congratulate Good Ventures for being the only people I know of in anything like their sphere of influence who are even talking about things like soil lead reduction.

So, a specific point here is that when we straighten out estimates of tractability, it appears that there are particularly exciting opportunities for **effective unilateral altruism in developed nations**. A more general point here is that Good Venture's rankings are highly sensitive to assumptions about the tractability of initiatives. This is worthy of additional attention, and is further reason why the full analyses of Good Ventures should be made public, if nothing else so that fresh eyes can apply a stress test to their calculations.¹⁹

7. Reliable Domestic Unilateral Altruism as a Competing Answer to the Big Question: A New Place to Look for Effective Altruism

Again, the problem we started with is that even though small charities like GiveDirectly arguably do huge amounts of good with several tens of millions of dollars in budget, they do not provide a good outlet for enormous donations, primarily because their operations cannot be effectively scaled up in the way that would be required to maintain anything like the same effect per donation.

In light of the discussion in the preceding sections, I propose that the following answer to the Big Question of how LARGE donations can do the most good is an important alternative to the answer given by effective altruists (which, again, was Answer 3):

¹⁷ See Anna Aizer, Janet Currie, Peter Simon, and Patrick Vivier. 2015. Lead Exposure and Racial Disparities in Test Scores, ms; for an overview, see Janet Currie's presentation, "Lead Exposure and the Black-White Test Score Gap", online at: <https://www.youtube.com/watch?v=dIW1KyMO7EU>.

¹⁸ See, for example, Nicholas Muller and Robert Mendelsohn. 2012. *Using Marginal Damages in Environmental Policy: A Study of Air Pollution in the United States*, for some discussion.

¹⁹ Thanks to Emily Clough for suggesting that the point here could be conceptualized as a missed opportunity for stress testing.

Answer 4: LARGE investments have the highest return when they are invested in boring low-hanging fruit, but of a different kind than has the highest return for SMALL investments. In short, LARGE investments have the highest return when invested in things that even at extremely large scale are boring sure things. These sure things are in domains where:

- (a) significant welfare gains are demonstrated as possible via existent rigorous economic analysis,
- (b) these gains are scalable in response to very large investments (even > \$1billion),
- (c) the evidence indicates that these gains will very likely be achieved even at this scale (in some cases by successfully influencing regulators who have existing unilateral regulatory authority),
- (d) will not be achieved without these philanthropic investments,
- (e) these gains are infeasible to achieve without large lump sum investment due to coordination problems (eg because there is no existing charity doing this work), and where ideally
- (f) these initiatives are not on the radar screen of philanthropists,²⁰ and
- (g) especially where these gains would disproportionately benefit the poorest members of society.

As noted above, some plausible causes that satisfy Answer 4 include pollution reduction along various dimensions where EPA authority already exists for that abatement, soil lead abatement, and other causes related to the intergenerational transmission of inequality described above.

Answer 4 seems likely to recommend different charities than Answer 3. One argument for this that I've in effect already made is that initiatives that meet criteria (a)-(g) present more reliable opportunities for effective altruism than those that are judged to be 'uncrowded windows of opportunity' by experts in philanthropy. As just one additional argument in favor of this, suppose that experts in philanthropy identify some alternative x to an investment that meets (a)-(g) as an 'uncrowded window of opportunity'; then x will tend to become crowded in philanthropic investment in the near future, and thus there is little expectation of positive marginal effect from investment in x. (Possible exception: if there is some barrier to philanthropists investing in this stuff; ie it is not sexy, or perceived to be morally wrong to invest there.) In contrast, opportunities that meet criteria (a)-(g) are very reliable (but boring) ways of achieving better returns in areas that would otherwise receive little or no investment even in the longer term.

One practical upshot suggested by Answer 4, in addition to a shift in criteria for investment, is that philanthropic advice aimed at LARGE donors should be based more directly on expert advice from applied economics and policy, where experts have special insight into investments meeting criteria (a)-(g), rather than based on judgments from the philanthropy community, and that effective altruism might be more often focused on domestic investment than it might appear on an a priori or even an initial a posteriori basis.

²⁰ Or: either unknown to philanthropy community, or there is a barrier to their entry – e.g. an initiative like kidney markets in which there is some moral repugnance associated with the initiative, or perhaps something in the charter of other foundations that prohibits funding.

One consequence of clarifying the structural properties of this kind of promising low-hanging fruit is that doing so also gives us some ideas for who to ask about where to find these opportunities: namely, we should not generally rely on philanthropy experts for insight into this, but should instead start with health and welfare economists and the like, with perhaps expertise from others who have insight into how best to design unilateral initiatives, including in some cases law school professors and the like when unlocking regulatory powers is essential.

At the very least, even if Answer 4 is ultimately wrong, by clarifying the structural properties of a particularly promising kind of low-hanging fruit, Answer 4 gives us an interesting place to look for promising investments from an effective altruism point of view, where at least some of these are likely to turn out to be a surprisingly good bet from an effective altruism perspective. This is, I take it, the practically important point in this neighborhood that we should have the most confidence is correct.

8. Contrarian Argument about Effective Altruism and Aid to the Global Poor

If Answer 4 were correct, then there would be an important contrarian argument about what Effective Altruism implies about aid to the global poor:

Contrarian Argument from Effective Altruism Against Aid to the Global Poor:

1. The best LARGE investments are in ‘boring low hanging fruit’ of the particular kind described in Answer 4.
2. Boring low hanging fruit of that kind tends to exist only in developed nations.
3. So, counterintuitively, the best LARGE investments tend to be in developed nations, even though there are far greater potential welfare gains in developing nations. This is true for a number of reasons, including a lack of governance and appropriate institutions that prevents effective small investments from being scaled up quickly to effective large investments with a similar return per dollar invested.

As the preceding discussion shows, there’s a surprising amount to be said in favor of this argument.

However, I don’t think this argument is necessarily sound, even if everything I’ve said up to this point is true. That’s because I think everything I’ve said so far is consistent with the truth of yet another possible answer to the Big Question of what LARGE investments do the most good:

Answer 5: LARGE investments have the highest return when they are invested in some of the same boring low hanging fruit that has the highest return for SMALL investments – except that LARGE investments in these initiatives do the most good when they are ‘metainvestments’ in building the necessary infrastructure and doing the necessary research to scale up these initiatives optimally, and so on. These investment opportunities are mostly in developing nations.

Alternatives to Answer 5 are also certainly possible.

What should we ultimately make of all this? Is Answer 4 correct, or Answer 5, or something else?

I don't have any confident view. Among other things, we need more information from experts on the relevant empirical issues that arise. (Another bit of unsolicited advice: effective altruists should stop alienating these experts, and instead bring them into the fold along with their vastly superior expertise on many specific issues.)

In what follows I argue that to answer the Big Question better metrics and more quantitative analysis seems necessary than is incorporated in Good Ventures's current analyses. Without that, and in the absence of future evidence, we should worry that the best LARGE investments are in boring low hanging fruit of the domestic kind, such a soil lead abatement and advocacy for tighter air pollution controls in our rich nations, even though there are far greater potential welfare gains in developing nations.

9. General Objections to the Effective Altruism's Standard Use of Tractability, Crowdedness, and Welfare Potential as Primary Metrics – and Philanthropy Dose Response Functions as an Alternative

Let me now set aside particular issues such as lead in soils, and just focus on a few conceptual questions about Good Venture's basic framework for making evaluations that we've been working with the whole time here. To do that, let's just look at their top ranked domestic initiatives.

E(Utility)	Priority	Cause	Welfare Potential	W	Tractability	T	Crowdedness	C
120	1	Criminal justice reform	Moderate	500	Unusual window of	80%	Moderate	30%
112	1	Macroeconomic stabilization	High	800	Highly uncertain	20%	Low	70%
100	2	Labor mobility	Very high	1000	Appears low	10%	Very low	100%
100	2	Land use reform	Moderate	500	Highly uncertain	20%	Very low	100%
70	2	Factory farming	Depends heavily on	500	Highly uncertain	20%	Low	70%
112	2	Marijuana policy	Low-moderate	200	Unusual window of	80%	Low	70%

Now I'm going to raise some objections to their basic framework here, but in doing this I see myself as very much a friend of Good Ventures, rather than any kind of a foe. In this way, one might see my comments in what follows as analogous to the sabermetrician followers of Bill James, who love baseball so much that they want to spend their free time trying to identify slightly better metrics that can be used by baseball teams to select players. Like any randomly selected advice from these folks, there's the risk that what I'll say is amateurish and wrong, or at least irritating and unhelpful in practice.²¹ Nonetheless, like these fans, I want my favorite team to win so much that I can't resist the temptation to offer the unsolicited advice that you're going to see in what follows. And I actually think that, as in the baseball case, it is actually pretty plausible that some crowdsourced "moneyball" can make these metrics much better.

So, with that warning, the next argument I want to make here is that there are a number of counterexamples to this framework (which again is endorsed by most effective altruists), which show that an initiative can be a very poor investment in terms of expected utility even if it does well along all three dimensions of welfare potential, tractability, and crowdedness, and vice versa.

²¹ But at least in the case of the baseball sabermetricians, their collective discussions over decades have definitely improved baseball. GiveWell and Good Ventures, with their transparency and blog posts can be seen as fostering a similarly welcoming environment for amateurs to contribute to their projects.

After presenting these counterexamples, I then explain how they point the way toward what I take to be some better metrics to replace the ones on Good Ventures’s spreadsheet, and also point the way toward some heuristics that might be useful in trying to quantify more precisely these better metrics.

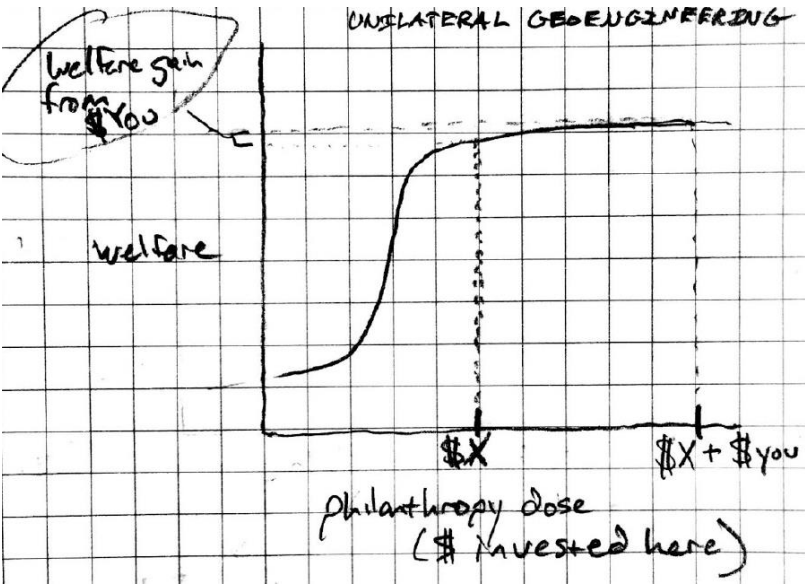
So, merely as an illustrative example, take the case of unilateral geoengineering by injecting sulfates into the earth’s stratosphere, which would counteract the general phenomenon of global warming. For my purposes here, it doesn’t matter what you think about this idea, I just want to use this to tell a story that makes a conceptual point based on a fictionalized story about this kind of intervention. So, in this story, suppose that absolutely no one is funding this except for Bill Gates, but that the technology has progressed to the point where it would only cost about \$100 million to actually deploy it, and a vulnerable nation with a good air force has agreed to allow Gates to build the infrastructure necessary to deploy it on its soil. Now let’s also assume, merely for the sake of a useful example, that a cautious amount of unilateral geoengineering should be expected to have great benefits on balance, primarily because it will give us very valuable information about how larger-scale geoengineering would actually play out in practice, and it would do this with little downside risk, whereas any more than a cautious amount of geoengineering would very quickly stop having additional benefits because of the downside risk and diminishing returns for global temperature.²²

Given this purely hypothetical stipulation of the case, I claim that it follows that in the sense intended on the Good Ventures spreadsheet, the Welfare Potential is at least moderate, the Tractability is good (because we are assuming it can be done unilaterally on a willing nation’s soil), and the Crowdedness is low because no one other than Bill Gates is currently investing in it:

E(Utility)	Priority	Cause	Welfare Potential	W	Tractability	T	Crowdedness	C
120	1	Criminal justice reform	Moderate	500	Unusual window of	80%	Moderate	30%
112	1	Macroeconomic stabilization	High	800	Highly uncertain	20%	Low	70%
100	2	Labor mobility	Very high	1000	Appears low	10%	Very low	100%
100	2	Land use reform	Moderate	500	Highly uncertain	20%	Very low	100%
70	2	Factory farming	Depends heavily on	500	Highly uncertain	20%	Low	70%
112	2	Marijuana policy	Low-moderate	200	Unusual window of	80%	Low	70%
280		Imaginary Geoengineering	Moderate	500	Good	80%	Low	70%

Now suppose we learn that Gates has actually paid the money to the contractors to in fact deploy a cautious amount of geoengineering in this way. As far as I can tell, this doesn’t change anything along the relevant dimensions of evaluation. Nonetheless, it then follows that any amount of additional investment by us in this domain will accomplish basically no good. That’s because the relationship

between philanthropic investments and welfare gains has something like this structure:



Here, we can imagine \$X dollars is the large investment that Gates has already made that is sufficient to capture the initial gains from a cautious amount of geoengineering. And it was stipulated that additional geoengineering beyond that point would have no positive effect. So, in light of this information, we should be confident in this case that the expected utility of an additional LARGE investment by us is basically zero, even though it would be an investment in a domain with high welfare potential, good tractability, and low crowdedness.²³

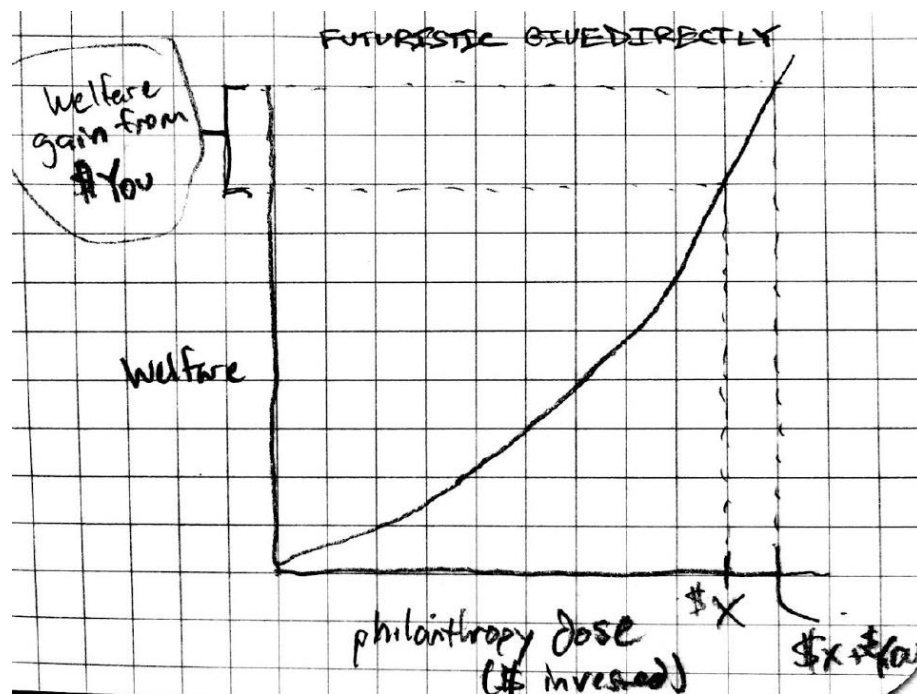
I propose to call this kind of graph a ‘philanthropy dose response function’, and I propose that construction of this kind of graph might at times be a useful heuristic for thinking about what welfare effect you should expect an additional investment in a domain to have.

Although the example I just gave is purely hypothetical, I think a similar problem might arise for some of Good Ventures’s top rated charities. To illustrate why, let me turn to a less imaginary example. So, consider the point in time five years ago, and suppose we are then evaluating the possibility of investing a LARGE amount in the cause of Homosexual Rights. Just as with the geoengineering example, it seems that there is a moderate level of potential welfare gain, low crowdedness, and as with criminal justice reform and marijuana policy now, it was true then that there was an ‘unusual window of opportunity’ in the tractability column. So, the actual rating along each dimension then would have been the same as we imagined in our hypothetical geoengineering example. Nonetheless, a structurally similar objection would have applied to investing more in homosexual rights five years ago, which we can see by noting that the philanthropy dose response function would have indicated that additional investments in homosexual rights then would have made essentially zero expected difference, in this case for a different reason: because although achieving homosexual rights was tractable then, that was a result of ongoing social change that was basically *independent* of philanthropy, and thus large philanthropic investments would not have made a difference in the odds of homosexual rights being achieved.

So, with this example in mind, I wonder why we should think that similar remarks do not apply to criminal justice reform and marijuana policy reform? I’m not suggesting that these cases are completely analogous in all interesting respects, but rather I’m merely highlighting the question of how much we should think that the probability of achieving these reforms depends on additional philanthropy, or whether instead the ‘unusual window of opportunity’ here and the prospects for achieving reform are mostly *independent* of philanthropy. If they are mostly independent, then this undermines the reason for thinking that it would be good to make additional investments in these causes, because it suggests that additional investments should not be expected to make a difference, just as in the imaginary geoengineering example. And again, this ‘no chance of making a difference’ phenomenon can realistically be present for a variety of reasons even when a cause does very well along the dimensions of welfare potential, tractability, and crowdedness – and there is some reason to worry that this phenomenon might actually present in connection with some of Good Ventures top ranked charities.

²³ Suppose you quickly build a station and ‘beat him to it’, and so your investment causes the welfare gain to be captured. Nonetheless, if his investment is already fixed, your investment is a waste, because it does not actually increase the amount of welfare with respect to global warming than if you had invested your money elsewhere, and it has the cost of not capturing the welfare elsewhere. So, it does no good, and it does bad by wasting resources.

As a final example, let me return to fantasy and ask you to imagine that 10 years in the future, givedirectly is now easily scalable, which we might imagine is because of background technological changes and institutional development, which makes it the case that we can now easily sign up the world's poorest people and deliver them money directly without risk of fraud or malfeasance by others. As a result of this, it is easy to imagine that now all kinds of money is flowing into this Futuristic GiveDirectly, from basically every charity that cares about the world's poor. As a result, Futuristic GiveDirectly would then score very high on the crowdedness metric. Nonetheless, all of this additional crowding could be correlated with increased emergence of standards as well as increased investments in infrastructure and thus increased competition between those providing services to GiveDirectly, which could actually increase the efficiency of the charity. More generally, high crowdedness should not count against additional donations to a charity insofar as that charity has a linear or convex philanthropy dose response function, as illustrated by this graph:



Here, the amount of money coming in implies that we are way out on the right hand side of the dose response function at point $\$X$. But being way out on the right hand side of *this* function does not entail that there aren't large additional gains to be had on the margin, or even that marginal utility is decreasing.

So, the upshot is that it not true that if a cause does well along dimensions of welfare potential, tractability, and crowdedness that additional investment in that cause then has high expected utility; on the contrary, in many realistic cases the expected utility of an additional investment will nevertheless be near zero. Similarly, the fact that a cause does badly along those dimensions is perfectly consistent with it being the best available cause. More explicitly estimating the philanthropy dose response function, and estimating the location of the margin on that function, could be a useful heuristic to use in trying to figure out what we should think about this issue in connection with particular cases. This heuristic can

also be used to estimate the value of what I take to be some better metrics to use in these evaluations, which I'll get to in a moment.

But first, here's my diagnosis of the problem with the existing metrics, focusing on the key factors in existing effective altruism analyses articulated above:

E(Utility)	Cause	Welfare Potential	Tractability	Crowdedness
-------------------	--------------	--------------------------	---------------------	--------------------

The problem as I see it is that these metrics have a particular meaning, where the meanings are basically this: welfare potential = potential welfare gains, tractability = the likelihood these gains can be captured, and crowdedness = the number of other philanthropy players trying to capture these gains (as a proxy for the percentage of these gains that would be captured by us conditional on gains being captured). More could be said to illustrate this, but I think it is now fairly clear that they don't reliably have the implications for the expected utility of additional investment in domains that they are taken to have by Good Ventures. As we've now seen, one key reason why these metrics fail is that a high likelihood that gains in a domain can be captured in some way or other does not imply that there is any substantial likelihood that adding more philanthropy to that domain will improve the chances that those gains will be captured – instead, it might be that the likelihood of these gains being captured is basically independent of philanthropy.

10. Better Metrics for Effective Altruism and the Big Question?

So, here's an initial crude proposal for some better metrics to use:

E(U)/\$	Cause	Welfare Gain Per \$ of LARGE Investment Conditional on that Investment Making a Difference	W D	Likelihood of an Additional LARGE Investment Making a Difference	Pr(D)
	Criminal justice reform				

Given the sort of rough estimates that Good Ventures seems comfortable making, I don't see why they wouldn't be comfortable making similar qualitative estimates of these metrics, and unlike their own metrics, these new metrics really would have the implications for expected utility that they ultimately want their evaluations to track.

So, let's return to the top contenders that we've discussed so far, and let's just begin by filling out 'Welfare Gain Per \$ of Large Investment Conditional on the Investment Making a Difference' by using Good Ventures's ratings of welfare potential, along with the same quantitative function from before for representing that as a cardinal ranking:

E(U)/\$	Cause	Welfare Gain Per \$ of LARGE Investment Conditional on that Investment Making a Difference	W D	Likelihood of an Additional LARGE Investment Making a Difference	Pr(D)
	Criminal justice reform	Moderate	500		
	Macroeconomic stabilization	High	800		
	Labor mobility	Very High	1000		
	Land use reform	Moderate	500		
	Factory farming	Depends heavily on val	500		
	Marijuana policy	Low-moderate	200		
	Soil Lead reduction	Moderate	500		
	EPA Lower Pollution	Moderate	500		
	WIC Supplement	Low-moderate	200		

Now, let me just, in a hand-wavy way, add what I take to be the most charitable estimate of the likelihood of making a difference that I'm comfortable making for Good Ventures's top rated causes, along with a similar estimate of the probability of making a difference for the 'intergenerational transmission of inequality' causes that we spent so much time discussing above:

E(U)/\$	Cause	Welfare Gain Per \$ of LARGE Investment Conditional on that Investment Making a Difference	W D	Likelihood of an Additional LARGE Investment Making a Difference	Pr(D)
5	Criminal justice reform	Moderate	500	Very Low	1%
8	Macroeconomic stabilization	High	800	Very Low	1%
10	Labor mobility	Very High	1000	Very Low	1%
25	Land use reform	Moderate	500	Low	5%
25	Factory farming	Depends heavily on val	500	Low	5%
10	Marijuana policy	Low-moderate	200	Low	5%
450	Soil Lead reduction	Moderate	500	Nearly Certain	90%
75	EPA Lower Pollution	Moderate	500	Moderate	15%
180	WIC Supplement	Low-moderate	200	Nearly Certain	90%

Now, I don't put *any* stock in this particular chart for reasons I'll get into in a minute, but I do think it makes the basic point that when we move to better metrics for actually estimating the expected utility of additional contributions to these causes, I suspect that this is going to move things away from thinking that Good Ventures's top rated charities are the best investment, and toward thinking that more **boring investments in fairly sure things** like pollution mitigation and early childhood health are better investments – where those latter investments, disturbingly, might only be available in developed nations.

At the same time, I don't any faith in this particular chart, because I don't think that these welfare ratings really make any sense in connection with the new metrics I've proposed, once you see that in order to make sense they have to be estimates of welfare gain per \$ of a very large investment conditional on that investment making a difference – and my intuition is that these numbers really don't adequately represent that, and in fact are unfairly skewed against Good Ventures's top causes.

So, in what is by far the hand-waviest thing I'm going to do here, let me just give my non-expert-but-honest-commentator intuitive gut feeling about my best guess about what the numbers in the welfare column might look like:

E(U)/\$	Cause	Welfare Gain Per \$ of LARGE Investment Conditional on that Investment Making a Difference	W/D	Likelihood of an Additional LARGE Investment Making a Difference	Pr(D)
0.2	Criminal justice reform		20	Very Low	1%
0.5	Macroeconomic stabilization policy		50	Very Low	1%
0.75	Labor mobility		75	Very Low	1%
0.15	Land use reform		3	Low	5%
0.25	Factory farming		5	Low	5%
0.25	Marijuana policy		5	Low	5%
1.8	Soil Lead reduction		2	Nearly Certain	90%
1.5	EPA Lower Pollution		10	Moderate	15%
1.8	WIC Supplement		2	Nearly Certain	90%

Now, there's no reason for you to take these numbers very seriously. But, I will say that there was no fine-tuning of anything here, except in the direction of trying to be overly charitable to Good Ventures's favored causes. I started with an estimate that I think makes sense of how much good additional cash transfers to the bottom three charities would do, in the 'real numbers' sense of e.g. doing \$2 of good for every \$1 invested in the sort of LARGE investment described above conditional on the intervention having the hoped effect, and then I tried to estimate the other welfare facts from there, trying to be overly charitable to Good Venture's favorite causes. Then I estimated the probabilities that the intervention would actually have the hoped effect. And then I let the expected utility estimates fall where they may.

I like this more explicitly quantitative approach for reasons discussed earlier, because if we are going to commit ourselves to quantitative claims anyway, let's make them as clearly as possible, and make them in a format that the top experts can readily engage with and review. I take it that the new metrics I'm using here have at least that virtue.

It may be useful to note that if you cut in half all of the probabilities for the highlighted causes, they still do better than the others. Alternatively, if you double the welfare estimates for the non-highlighted charities, they still don't do as well as the highlighted ones. You really have to do something like both of those things to get rough parity between the highlighted and non-highlighted causes here.

But again, apart from the metrics themselves, and the methodology of explicit real numbers quantification – which I suspect would be genuine improvements – there’s no reason to put faith in the actual numbers here. So, let’s not focus on them, but focus rather on, again, the different possible answer that they suggest to the Big Question of how LARGE donations can do the most good – namely, via particularly exciting opportunities for **effective unilateral altruism in developed nations**.

As a final note, I should stress that my suggestions here are merely for a few crude improvements over the existing metrics of effective altruism. For an economic analysis that suggests more precise metrics that could be used, in other work²⁴ my coauthor Dean Spears and I analyze the marginal effect of a contribution to charity using the following equation:

$$\left(\frac{\Delta \text{Lives Saved}}{\Delta \text{Donation}}\right) = \left(\frac{\Delta \text{Budget}}{\Delta \text{Donation}}\right) * \left(\frac{\Delta \text{Activity}}{\Delta \text{Budget}}\right) * \left(\frac{\Delta \text{Lives Saved}}{\Delta \text{Activity}}\right)$$

In that paper, we offer a number of suggestions for moving effective altruism toward a more explicitly marginalist orientation and metrics, which could be focused on such an equation. I endorse those suggestions here as well as a guide to what evaluations should ultimately be tracking.

[END]

²⁴ Mark Budolfson and Dean Spears, “Effective Altruism, Marginal Impact, and Fundraising: Weak Links in Effective Altruism’s Chain”.