Shallow Report on Asteroids
October 2022
Joel Tan, CEARCH

Summary

Factoring in the expected benefits of preventing asteroid impact events (i.e. fewer deaths and injuries) as well as the tractability of lobbying for asteroid defence, I find that the marginal expected value of such asteroid defence lobbying to be **1,352 DALYs per USD 100,000**, which is around 2x as cost-effective as giving to a GiveWell top charity.

Discussion:

- The results surprised me – I came in expecting extremely low marginal expected value to working on this cause, given low neglectedness (c.f. DART), low risk of occurrence (c.f. fairly low per annum death rates compared to the likes of malaria), and middling tractability (i.e. effective but expensive interventions available).
- I do not model the expected costs of dual use (i.e. deflection technology potentially being used as a weapon to kill or injure vast numbers of people), as such a possibility, slight or otherwise, does not truly increase existential risk. (a) Firstly, there is no reason for the great powers to ever deploy or develop planet-killing kinetic bombardment capabilities (i.e. ones capable of causing catastrophic global winter or an immediate extinction-level event) – these cannot be used without destroying yourself, unlike nuclear weapons where there is the theoretical possibility of a successful first strike. (b) Secondly, while the great powers may see military use for smaller scale orbital bombardment weapons (i.e. ones capable of causing sub-global or Tunguska-like asteroid events), these are only as destructive as nuclear weapons and similarly cannot be used without risking nuclear retaliation. This means that such weapons are functionally identical to any other missile in a nuclear arsenal, whose existence (given nuclear weapons already exist) does not – at the margin – increase the risk of death and injury for humanity, whether from direct kinetic exchange or from climatic shifts due to full-scale WMD exchange.
- For interventions, funding additional search telescopes (whether ground-based or space-based) isn't considered because existing assets, augmented by the soon-online Vera C. Rubin Observatory and NEO Surveyor, will probably identify all the >140m near-earth objects – which pose 99% of the risk – by 2028. Any new telescopes would literally not be built or launched before we succeeded with current assets.
- Encouraging results are driven by high tractability, which is itself a function of (a) a large fraction of the problem being solved by the intervention, and (b) heavily discounting the counterfactual value of US government funding relative to EA funding.
- Very sensitive to costing estimates

- If further research is done, expert advice on tractability – both on the political aspect (lobbying success) and on the technical aspects (ATLAS and deflection efforts) – would be very valuable.

Expected Benefit: Averting Asteroid Impact Deaths

The primary expected benefit of preventing asteroid impact is averting deaths. This benefit is modelled in the following way.

Moral Weights: I take the value of averting one death to be **29.3 DALYs**. This is calculated as a function of (a) a human's full healthy life expectancy of 63.69, (b) a minor age-based philosophical discount, and (c) assuming we save someone of the median age in the relevant population. For more details, refer to CEARCH's evaluative framework.

Scale: To calculate the deaths from asteroid impact, I consider four scenarios: (1) deaths from a Tunguska-like asteroid event (i.e. the impactor being 50-300 m in diameter); (2) deaths from a large sub-global asteroid event (i.e. the impactor being 0.3-1.5 km in diameter); (3) deaths from a global asteroid event (i.e. the impactor being >1.5 km in diameter); and (4) deaths from rare K/T scale asteroid event (i.e. the impactor being >10 km in diameter). Moreover, to calculate deaths within each scenario, I use three different estimates (i.e. Morrison & Chapman, Stokes et al, as well as the National Research Council). I adjust for population growth since the estimate was made, and take a weighted average of the three estimates, penalizing the Chapman & Morrison estimate for not taking into account tsunami effects.

Overall, this yields the following deaths per event:
- Tunguska-like asteroid event: **145,000 deaths**.
- Large sub-global asteroid event: **2.06 million deaths**.
- Global asteroid event: **3.4 billion deaths**.
- Rare K/T scale asteroid event: **6.3 billion deaths**.

Persistence: If we managed to eliminate asteroid risk, the per annum benefits would persist over time and would hence need to be summed up across various years, but there are a couple of important discounts that we need to implement.

Firstly, we must discount for the probability of the solution not persisting (i.e. an effectual planetary defence being dismantled). Assuming our intervention is lobbying Congress for more money so that NASA's Planetary Defense Coordination Office can better destroy any asteroids that will potentially hit and harm earth, the risk we face is that the PDCO is defunded further down the line. I take this to be the risk that NASA suffers budget cuts – as calculated from historical trends (26 years in which cuts occurred out of 63 years surveyed)– multiplied by the risk that the PDCO in particular then suffers significant budget cuts which makes it unable to carry out its core mission (n.b. I assign a further 1% chance to this, on the basis that the PDCO will likely be prioritized due to its obvious importance and it being only around 0.7% of NASA's budget in any case. Overall, this gives a reversal rate of **0.41%**.

Secondly, we must discount for the proportion of the solution being counterfactually solved (i.e. planetary defence being conducted without additional intervention). Here, I look at three specific considerations.

- (a) Existing success of NASA, as supported by by various US government agencies, non-profits, corporations and foreign countries: Of the around 25000 >140m near-earth objects (NEOs) that pose 99% of the risk to humanity, the NASA-led effort has identified around 9951 as of early October 2022 (i.e. around 40%), and if any turns out to be a threat, we would have the time to prepare countermeasures and subsequently destroy or deflect the asteroid/comet with a 97.76% chance of success (refer to the section on tractability for the calculation of this figure). Overall, this means that NASA is on track to prevent **38.71%** of any harm.

- (b) Expected future success of the NASA-led effort in mitigating asteroid risk: For the remaining >140m NEOs that pose 99% of the risk to humanity, the current NASA-led effort is discovering them at a rate of around 500 a year (i.e. 2% of these NEOs).

  Meanwhile, the Vera C. Rubin Observatory (i.e. a ground-based shared LSST) is scheduled to come online in 2024. Per the NASA 2007 NEOSDAA report examining whether a ground-based shared LSST be built, and considering the last 5 years' track record of 350 potentially hazardous objects (PHOs) found per annum, business as usual from a 2007 baseline would have seen 7,824 PHOs (31.3%) found by 2020, while a ground-based shared LSST (originally supposed to starting operation in 2014) would have reached 75% completion (i.e. 18,750 PHOs found) by 2020 – which means the LSST is capable of finding an additional 1,561 PHOs (i.e. 6.2%) per annum.

  Further, the NEO Surveyor (i.e. a space-based 0.5m IR telescope @ L1) will be launched in 2026. The NASA 2007 NEOSDAA report similarly considered this L1 satellite as a option for searching for asteroids and comets, and it was expected that with it starting operations in 2013, we would reach 85% completion (i.e. 21,250 PHOs found) by 2020 – accordingly, the L1 satellite can find an additional 1,678 PHOs (6.7%) per annum.

  Both the LSST and the L1 satellite were delayed, and since they will only start operations in the future, their impact are partially discounted.

  Overall, with the above detection rates factored in and taking into account a subsequent asteroid destruction/deflation rate of 97.76%, this yields a per annum discount rate of around **12.92%** from this cause area becoming increasingly less neglected going forward.

- (c) In terms of structural factors – while trends might theoretically make concerted planetary defence efforts more likely (e.g. economic growth making us more able to fund planetary defence efforts, or cultural change moving us towards post-materialistic values which in turn cause us to be worried about the long term future rather than our own short term interests), this will be a vanishingly weak effect, and I do not bother modelling it (i.e. a **0%** discount).

Thirdly, I apply a standard existential risk discount. Here, I factor in the probability of total nuclear annihilation, since the benefits of saving people from asteroid impact in one year is nullified if they had already died in a previous year. For the exact risk of total nuclear annihilation, I take it to be one magnitude lower than the risk of nuclear war itself, since nuclear war may not kill everyone. For the probability of nuclear war, I use the various estimates on the probability of nuclear war per annum collated by Luisa, but with accidental nuclear war factored in, and then calculate a weighted average that significantly favours the superforecasters. The reason for this is that (a) the estimate of the probability of intentional nuclear war based on historical frequency is likely biased upwards due to historical use being in a MAD-free context; (b) the probability of accidental nuclear war based on historical close calls is highly uncertain due to the difficulty of translating close calls to actual probabilities of eventual launch; and (c) experts are notoriously bad at long-range forecasts, relative to superforecasters. Meanwhile, I do not take into account other existential risks like supervolcano eruption and asteroid impact, since the chances of those occurring at all is very marginal per Denkenberger & Pearce, let alone the chances of such events killing everyone and not just most people. Overall, therefore, I treat the general existential risk discount to be just the risk of nuclear war but adjusted a magnitude down (i.e. **0.07%**)

Fourthly and finally, I apply a broad uncertainty discount of **0.1%** to take into account the fact that there is a non-zero chance that in the future, the benefits or costs do not persist for factors we do not and cannot identify in the present (e.g. actors directing resources to solve the problem when none are currently doing so).

Value of Outcome: Overall, the raw perpetual value of averting asteroid impact deaths is:

- Tunguska-like asteroid event: **1.94 * 10^7 DALYs**.
- Large sub-global asteroid event: **2.76 * 10^8 DALYs**.
- Global asteroid event: **4.55 * 10^{11} DALYs**.
- Rare K/T scale asteroid event: **8.42 * 10^{11} DALYs**.

Probability of Occurrence: For the probability of a Tunguska-like asteroid event, I look at three estimates – Chapman & Morrison, Stokes et al, and the National Research Council. The NRC figures obtained are relative approximations, but Stokes et al is particularly detailed in listing the impact probabilities of potentially hazardous objects (i.e. near-earth asteroids and extinct short-period comets with a minimum orbit intersection distance of <0.05 astronomical units from the Earth's orbit) – for diameter 0.0313 km, the frequency per year is 3.17E-03; for

0.0394 km, 1.84E-03; for 0.0496 km, 1.07E-03; for 0.0625 km, 6.19E-04; for 0.0787 km, 3.60E-04; for 0.0992 km, 2.09E-04; for 0.125 km, 1.21E-04; for 0.157 km, 7.03E-05; for 0.198 km, 4.08E-05; for 0.25 km, 2.37E-05; for 0.315 km, 1.38E-05; for 0.397 km, 7.99E-06; for 0.5 km, 4.64E-06; for 0.63 km, 2.69E-06; for 0.794 km, 1.56E-06; for 1 km, 9.07E-07; for 1.26 km, 5.26E-07; for 1.59 km, 3.06E-07; for 2 km, 1.77E-07; for 2.52 km, 1.03E-07; for 3.17 km, 5.98E-08; for 4 km, 3.47E-08; for 5.04 km, 2.01E-08; for 6.35 km, 1.17E-08; and for 8 km, 6.79E-09. For the 50-300 m bucket, therefore, the aggregate probability is around 2.53E-03. Aggregating this with the other two estimates, I arrive at a weighted average, wherein (a) I penalize the Stokes et al estimate (given the possibility of my making an error when aggregating the discrete probabilities, and (b) I penalize the NRC estimate given that the figures sourced from there are relative approximations. Overall, this yields **0.5%** probability per annum of a Tunguska-like event occurring.

For the probability of a large sub-global asteroid event, I again look at the Chapman & Morrison estimate, the Stokes et al estimate (as summed from the discrete probabilities, yielding an aggregate probability of around 3.24E-05 for the 0.3-1.5 km bucket) and NRC estimate (again an approximation). As before, I penalize Stokes et al and the NRC in the weighted average, for reasons highlighted above. Overall, we have a **0.004%** probability per annum of a large sub-global asteroid event occurring.

For the probability of a global asteroid event, I once more rely on Chapman & Morrison, Stokes et al (aggregate probability is around 7.19E-07), and NRC (once again an approximation), and perform the weighing in a similar fashion as above for similar reasons, to get a **0.0002%** probability per annum of a global asteroid event occurring.

Finally, for the probability of a rare K/T scale asteroid event, I shift to relying on Shoemaker, Wolfe & Shoemaker on top of Chapman & Morrison as well as the NRC. They all agree on the precise probability of a >10km asteroid hitting earth (**0.000001%** per annum), and the simple average is just that.

Expected Value: Overall, the expected value of averting asteroid impact deaths is **$9.36 * 10^5$ DALYs**.

Expected Benefit: Averting Asteroid Impact Injuries

Beyond killing people, the impact of an asteroid or comet hitting earth can cause injuries not amounting to death, the effects of which I also model.

Moral Weights: I take the value of averting a typical injury in nuclear war to be **5.88 DALYs**. This is calculated as a function of (a) the average disability weight for all injuries, (b) a minor age-based philosophical discount and (c) assuming we save someone of the median age in the relevant population. For more details, refer to CEARCH's evaluative framework.

Scale: As Stokes et al note, the blast damage from an asteroid's impact on land is similar in nature and scale to the damage from nuclear explosions, so I use the injury-to-fatalities ratio from a nuclear war context here. I consider three separate reference points – the Wellerstein et al estimate of NATO-Russia nuclear war, the Kristensen, Norris & McKinzie estimate of US-China nuclear war, and a US intelligence estimate of India-Pakistan nuclear war, yielding an average injury-to-fatalities ratio, which I then apply to the already-calculated fatalities per event category:

- Tunguska-like asteroid event: **108,000 injuries**.
- Large sub-global asteroid event: **1.9 million injuries**.
- Global asteroid event: **71.9 million injuries**.
- Rare K/T scale asteroid event: **248 million injuries**.

Persistence: The same discounts discussed in the section on asteroid deaths are applied here.

Value of Outcome: Overall, the raw perpetual value of averting asteroid impact injuries are as follows:

- Tunguska-like asteroid event: **$2.9 * 10^6$ DALYs**.
- Large sub-global asteroid event: **$5.1 * 10^7$ DALYs**.
- Global asteroid event: **$1.93 * 10^9$ DALYs**.
- Rare K/T scale asteroid event: **$6.65 * 10^9$ DALYs**.

Probability of Occurrence: The same probabilities discussed previously are applied here.

Expected Value: All in all, the expected value of averting asteroid impact injuries is **$1.7 * 10^4$ DALYs**.

Tractability

To mitigate the risk posed by asteroid impact, I consider the potential solution of lobbying for asteroid defence, the theory of change for which is as follows:

- Step 1: Lobby the US Government (i.e. White House and Congress) to fund a standby planetary defence rocket.
- Step 2: ATLAS performs last minute detection of larger asteroids not yet found, or of smaller asteroids that aren't being systematically catalogued
- Step 3: Use of standby planetary defence rocket to successfully destroy or deflect incoming asteroid.

Step 1: The main outstanding risk, given that the NASA-led effort is on track to catalogue all the >140m PHOs, is that either a >140m asteroid is on track to hit earth before the cataloging is complete, or that smaller asteroids slip through more generally, giving us not enough time to prepare mitigation efforts (n.b. five years of preparation would be needed for a mission to destroy or deflect an asteroid). Consequently, it would be useful to have a planetary defence rocket on standby – but whether that is achievable depends on our ability to successfully lobby the US Government accordingly.

To assess the probability of lobbying success, I take both an outside and inside view.

For the outside view, I consider three reference classes. First, there is the general success rate of Washington lobbyists (56%). Secondly, I consider past instances of budget increase for planetary defence at NASA. The logic here is that bureaucracies always push the legislature and chief executive for more money, so any actual subsequent increase is indicative of the latter being supportive. Summed across the years, we can hence get a sense of the probability in any particular year that the White House and Congress will be supportive and likely to agree to additional funding for any given purpose. And indeed, 92% of the time, planetary defence gets a year-on-year budget increase. Thirdly, I consider past instances of a budget increase at NASA more generally (same logic as above). Overall, NASA has enjoyed a budget increase in 33 out of 63 years surveyed (52%). In taking a weighted average of the three estimates, the planetary defence/NASA reference classes are penalized for relying on an uncertain chain of logic, and the NASA reference class is then penalized two more times again on top of that, firstly for relying partly on decades-old data; and secondly for just being less relevant than the planetary defence reference class

For the inside view, my reasoning is as follows. Congress is generally supportive of planetary defence, but the current Biden administration seems sceptical, and together with the significant cost of the standby planetary defence rocket – assuming similar figures to the Space Launch System – I estimate that chances of success are fairly low, perhaps 10%.

Putting the outside and inside views together, I give more weight to the outside view than the inside view, given that the inside view is subject to the usual worries about inferential uncertainty. This yields a 54% chance of success at this stage of the theory of change.

Step 2: Of course, even after a hypothetical planetary defence rocket is ready to go, it can only be launched to destroy or deflect an incoming asteroid if the latter object is detected in the first place. The Asteroid Terrestrial-impact Last Alert System (ATLAS) serves as our last line of defence (assuming other detection systems have failed to detect the incoming threat at longer ranges), but even ATLAS isn't foolproof – the smaller the object, the harder it is to detect. Assuming that 50-140m asteroids are the main remnant threat, I take an average of the probabilities of detection for 50m and 140m asteroids, yielding a 59% detection probability.

Step 3: Having detected the incoming threat, the planetary defence rocket would be launched for intercept. As for whether destruction or deflection is even conceptually possible – (a) in the impulsive category, the use of a nuclear device is the most effective means to deflect a potentially hazardous object (PHO); (b) meanwhile, non-nuclear kinetic impact alternatives (i.e. the DART method) are the most effective non-nuclear option, transferring 10-100 times less momentum than nuclear options for a fixed launch mass; and (c) for slow push techniques, they are useful in relatively rare cases (<1% of expected threat scenarios). The upshot of all this, in any case, is that asteroid deflection is certainly technically plausible given adequate funding.

As for the specific probability that an attempted intercept achieves success – I take an outside view to evaluate this. An inside view is not considered, insofar as I lack the technical expertise to make any such determination. For the outside view, the three reference classes I consider are as follows. Firstly, DART, which successfully altered Dimorphos's orbit around Didymos by 32 minutes, showing that if an Earth-threatening asteroid was discovered sufficiently early it could be deflected by kinetic methods (100% success rate from a sample size of 1). Secondly, I consider the rate of success for NASA's human spaceflight missions – 166 such missions were conducted, of which 3 ended in deadly failure (98% success rate). Thirdly, I consider the rate of success for SpaceX rocket launches (97% so far). I then take a weighted average, doubly penalizing DART because: (a) it has a far lower weight (relative to the payloads we'll likely need to deflect threatening asteroids), unlike NASA space shuttles or SpaceX rockets, the upshot of which is that more force must be generated for lift, which brings with it greater risks; and (b) DART offers but a sample size of 1. Overall, this suggests a 98% chance of success for a launch of a planetary defence rocket on a destruction/deflection mission.

Putting the estimates across all three steps in the theory of change together (i.e. the probability of successfully lobbying the US Government to fund a standby planetary defence rocket, the probability of ATLAS managing last minute detection of larger asteroids not yet found, or of smaller asteroids that aren't being systematically catalogued, and the probability of mitigation success), the intervention is expected to eliminate 32% of the remaining risk from asteroids.

Moving on, let us consider the costs to all this. On the one hand, we have the cost of hiring lobbyists to persuade the US government; I compute this as the average of firms' top-end spending on K-street lobbying in 2021 (around 21.7 million USD). On the other hand, we have the immense costs of building and operating the hypothetical planetary defence rocket. However, this cost has to be discounted, and fairly significantly: (a) for the probability that it is even incurred (not so, if we fail in the lobbying); and (b) the lower counterfactual cost of US government spending relative to EA funding going to top GiveWell charities or similar. The latter is in turn a function of diminishing marginal utility of income, higher US GDP per capita relative to the poor country average, and the top GiveWell health charity's cost-effectiveness relative to just giving cash to poor people, correcting for GiveWell potentially undervaluing life vs income. Overall, I estimate that the counterfactual cost of the standby rocket itself is only around 593,000 dollars. In total, this puts the total cost of the intervention at 22.3 million dollars.

Consequently, the proportion of the problem solved per additional USD 100,000 spent is around **0.00142**.

Marginal Expected Value of Lobbying for Asteroid Defence

Combining everything, the marginal expected value of lobbying for asteroid defence is **1,352 DALYs per USD 100,000 spent**, making this 2x as cost-effective as a GiveWell top charity.