How large is the value of humanity's future?

Kevin Kuruc and Gage Weston
(with help and input from the PWI team)

May 2023

Background Context: This report and accompanying Google Collab Notebook[1] makes progress on quantifying the value that the future could hold, conditional on surviving this century. In particular, we ask what can be conservatively claimed, in an effort to see how far one can get without relying on some of the extremely high value sci-fi-like scenarios discussed in longtermist circles (e.g., we rule out futures with large numbers of sentient digital minds). In other words: **we ask what the value of a purely biological future could be.** Should we expect the future to be many orders of magnitude more valuable than the present century even with the prospects of low-fertility and declining human populations? What if the tail scenarios are truncated to prevent fanatical considerations from dominating expected value calculations? We focus on these questions in order to study whether Effective Altruists might agree on a lower-bound for the value of preserving the future, even while potentially disagreeing about the importance of digital minds, tail outcomes, etc.[2] To this end, where applicable, we also hedge away from assumptions of large qualitative changes to human life that might be seen as controversial.

Method: We work within a total utilitarian framework where the well-being of potential people is treated as equivalent to already-existing people and break the analysis into three dimensions: Long-run average population sizes; long-run rates of extinction risk ('x-risk'); long-run average quality of life. We specify a full distribution of credences over different long-run outcomes for each, where "long-run" represents the average value over all future periods. Then—because the statistical properties of multiplying across these distributions are complicated in a way that precludes a simple analytical solution—we perform Monte Carlo simulations that draw from the respective distributions. The result is a distribution of outcomes from which the expected value, median value, non-fanatical-expected-value, etc. can be calculated.

Findings: The expected value of the future seems large in this framework (>10,000x the value generated this century). This is driven by the right-tail of outcomes, even in our attempt to be conservative. For example, the baseline credence distributions imply a median value of the future that is two orders of magnitude smaller than the expected value (≈350x the value of this century). To the extent one wants to truncate the tails to reduce the influence of the extreme outcomes, the adjusted-expected-value can be significantly smaller.

The underlying properties that drive these results are easy to understand: the possibilities for annual-population-size and centuries-survived have long right tails. Both variables are bounded below by zero, so uncertainty serves to drive up the expected value of each. Quality of life probably follows a similar pattern.[3] The product of these terms will then inherit an even more extreme tail: the futures with large populations, very high quality of life, and that are long-lasting generate a disproportionate amount of the expected value. Determining the value of the future then depends on how much weight one feels comfortable assigning to extremely good outcomes, even in this conservative, biological approach where we rule out the scenarios[4] typically associated with fanatacism.

[1] Available upon request; contact Gage Weston (gage@weston.co) or Kevin Kuruc (kevinkuruc@utexas.edu).

[2] Of course, other empirical facts are necessary for determining whether we ought to invest in reducing the probability of near-term existential catastrophe in hopes of attaining this value. As an obvious example, if it were impossible to change the probability that we go extinct, then it would not make sense to invest in doing so regardless of the value of the future.

[3] Though this variable differs in that it is conceptually possible to have quality of life that is negative.

[4] Here again we're ruling out futures with very large numbers of digital minds.

1 Long-Run Population, X-Risk, and Quality of Life: A significant chance of a low-value future

Here we discuss our views about the future of population sizes, rates of existential risk, and changes in quality of life. The median long-run population outcome we assume is less than 10 billion because global projections for fertility this century imply population decline to below this number in the medium-run. Furthermore, there are no known forces or incoming technologies that we feel confident will result in replacement rate fertility; the sum of social scientific facts points towards a desire for fewer children in a world of abundance. Re-entering a phase of persistent growth is correspondingly less likely. That said—and a theme of this write-up—our uncertainty over this gives rise to a right tail and an expected value that significantly exceeds the median and modal outcomes.

Regarding x-risk, we have less to say that is novel. The *time-of-perils* narrative—wherein x-risk is high now but predictably low in the future—pushes things in favor of longtermism by assumption. In the spirit of epistemic humility and conservatism, we do not put most of our probability mass on very safe futures. But, here too, the expected number of centuries significantly exceeds the median number of centuries survived. Quality of life seems less important: in the space of non-sci-fi futures we are interested in, orders of magnitude differences in quality of life seem less likely. We spend less time discussing our views on this variable and assume it will be within an order of magnitude of its current value with high probability.

1.1 Long-run average populations

Projecting the future of population sizes is vexing. The consensus among demographers is that by the end of this century global fertility rates will be below two, the rate we would need for each generation to replace itself (e.g., Raftery, 2021; UN, 2022; IMHE, 2022; IASSA, 2022). This would imply that the medium-term, at least, will contain a phase of population decline. Furthermore, to a first-approximation, the only societies across time-and-space in which the average woman has (had) two children have been poor and technologically immature.[5] If:

1. Fertility projections imply depopulation and
2. All observable modern societies have sub-replacement fertility then
3. It is extremely difficult to guess when, if ever, this decline would stop.

Below we expand on the drivers of low-fertility and discuss what sorts of developments are (not) likely to reverse this trend. We then put a rough probability distribution over future long-run population sizes.

1.1.1 A worrying possibility: Opportunity cost theory → a vanishing population

The background empirical trends that seem most important to us are the following: since the industrial revolution (1) living standards have gotten much *better* and (2) fertility rates have gotten much *lower*. Currently, the PWI working theory of low-fertility is that the high opportunity cost of time in a world with many more rewarding careers, leisure, and other opportunities has resulted in a decline in the *relative* value of having a child. When there are more good things to do with your time, having a kid appears to become relatively less attractive.

Quantitative research has shown that other common explanations for current low-fertility have a difficult time explaining its universality; very different socio-economic-cultural contexts have converged on similar behavior (e.g., Brazil, Sweden and the US all have estimated fertility rates of 1.66). If the key drivers were

[5] Modern Israel is one important exception to this. We return to the possibility that the fertility rates of the religious will reverse population decline later.

housing or educational costs, women's labor force participation, reduced intergenerational mobility, etc., it would not be the case that this phenomenon would arise in all places that have reached a modest level of economic development. Western Europe has a more robust social safety net and subsidized college; Japan has kept housing prices more manageable than the US; etc. And yet, all of these countries are characterized by low fertility rates. An increase in life opportunities for women, however, seems to be relatively universal in regions that have developed, all of which now have low fertility.

Taking this opportunity costs theory seriously leads to bleak predictions for future population sizes: as the world gets even better, we might expect fertility rates to get even lower. Where this stops is difficult to know. It either needs to be the case that some important social, policy or technological change occurs that makes parenting a much better/more attainable proposition, or that the quality of life becomes so low that having children returns to being an attractive way to spend time and resources.

In general then, we believe that the obstacles to achieving replacement rate fertility are often underestimated. The future may very well be one of substantial population declines before stabilization or growth. Consequently, our distribution of potential population sizes assigns greater weight to smaller populations than other longtermists might. Prior to presenting our detailed quantitative distribution, we will address counter-arguments that we have encountered for reasons to expect fertility to rebound, along with rationale supporting the more pessimistic outlook.

1.1.2 Forces that may (not) lead to significant fertility increases

Here we discuss—roughly ordered from less likely to more likely, in our view—developments that we've seen proposed for why fertility might rebound in the future. This is meant to broadly illustrate the case for pessimistic population outlooks before formally parameterizing a distribution of outcomes.

Space Settlement. Some longtermist back of the envelope calculations about future population sizes leverage the possibility of becoming multi-planetary. But this reasoning, to a first-order, relies on land or natural resources being the binding population constraint. Why else would we expect more habitable planets to result in more people, rather this merely resulting in more geographic dispersion?

To draw this out, consider our own planet. There is still plenty of room in regions much more hospitable than other planets (e.g., the western United States) where natural resources remain easily accessible. This makes us pessimistic that increasing the land and resources available to families will change fertility behavior. Furthermore, the individuals that prioritize having many children can much more easily achieve that without also emigrating to a new planet.[6] It's certainly not obvious that higher fertility individuals would be the ones choosing to settle space. In general, if space settlement occurs, it seems like it will be *downstream* of already large population sizes, not something that will help prevent depopulation.

Artificial Wombs (on their own). This conjecture is closer to the mark: this technology would reduce the opportunity costs of having children. Would it be enough to change fertility behavior? It doesn't seem so, at least on its own. Pregnancy is a minuscule fraction of the effort spent raising a child over their entire lifetime. While this sort of technology would also eliminate the age constraint on having children, already-existing assisted reproductive technologies have not importantly changed aggregate trends, even in places where it is more widely available and subsidized (Lazarri et al, 2021). The demand for newborns among women out of childbearing years appears too low to make a large difference. Later we discuss how artificial wombs *paired* with other technologies might change fertility dynamics. But we do not believe this technology is sufficient to significantly shift fertility behavior.

Governments increasing their financial commitment to combating low-fertility. The first reason we expect this to be difficult is current empirical evidence: present day governments in very low fertility settings that care about this issue and have made no important progress (e.g., Japan, South Korea, Hungary).

However, future governments will have more resources and/or incentives to solve this problem. For example, in a richer world, governments' pro-natalist budgets may be larger (in per capita terms). While

[6]As Elton John sings in *Rocket Man* "Mars ain't the kind of place to raise your kids. In fact, it's cold as hell."

true, the opportunity costs for potential parents will have risen in tandem, meaning the required pro-natalist payments to influence fertility will have also grown. We do not see a strong reason to believe governmental resources will grow systematically faster than the forces that make having children a relatively low-value activity.

Alongside additional fiscal resources, governmental incentives to enact such policies will plausibly increase once labor shortages and/or pension shortfalls become more acute. In such scenarios *current* resources will have also become more valuable. To be specific, if governments are responding to labor shortages they are doing so in a context where wages are high, which makes time spent away from work raising children more costly to parents. Furthermore, if governments were in general concerned about contemporaneous living standards or even long-run economic growth, it is very unlikely that fertility policy would be on the list of top policy prescriptions. For example, if there were pension shortfalls, we would expect excess governmental funds to go towards directly filling those gaps, not incentivizing fertility to indirectly solve the problem many years in the future. And insofar that population size matters for R&D, we do not believe anyone claims it is the most cost-effective way to increase economic growth.[7]

Heritable Fertility. Some sub-cultures have above replacement fertility (e.g., many religious groups). This might imply that as the (now-dominant) culture with below replacement fertility shrinks, the sub-cultures with above replacement fertility will become a larger share of the population. Once they are a sufficiently large fraction, the aggregate data will converge to the behavior of that group, by definition.

One way to immediately see that this reasoning relies on additional implicit assumptions is to look at the share of the Western world that identifies as religious. The (Western) world is becoming more secular, despite the religious sub-population having systematically higher fertility. In this case, *cultural retention* is the obvious culprit for the discrepancy. Indeed, the heritability theory, on its own, would *never* predict declines in aggregate fertility rates; it should always be true that higher fertility sub-populations become an ever-larger share of the total population (Collins and Page, 2019). Clearly, other factors over the last 200+ years have been much more important than heritability; we expect in the future much of the same to be true.[8]

Because low fertility appears to be the revealed preference across a very diverse array of modern cultures, we treat it as the likely default among future people with better lives than our own. This means that such a high-fertility sub-culture would need to permanently buck this trend, even as it spread across time and space. It is not impossible, but it is more speculative than the simple differential equations imply.[9]

Dramatic declines in the cost of parenting. So far, we've made a lot of progress automating the creation of goods, such that the relative costs of parenting and other "service" style activities have gone up. In a "post-work" // "post-scarcity" scenario, you could imagine there is a catch-up and we become relatively more productive at service style, labor intensive activities. If we become much better at parenting (either by automating many unpleasant tasks, finding communal living set-ups, or whatever else makes parenting less burdensome), that could plausibly raise fertility.

Change in Values and/or Social Pressure. Right now, there is social status attached to performing well in a career, for example, while very little seems to be attached to raising children. We could imagine a future where almost all economic tasks are automated, laying the groundwork for very different values or choices. If abundant leisure time leads people to search for more meaningful activities, perhaps raising children will be one of those.

This could be further bolstered by high-fertility becoming seen as pro-social, and therefore higher status than it is today. If worries about depopulation become widespread, or if our moral concerns expand to

[7] Even if it is the most effective way to promote very long-run growth—because, for example, you may be creating a whole line of descendants with fertility policy—a myopic government facing political pressure is unlikely to optimize for the very long run.

[8] An important exception to that are Orthodox sects in Israel that have grown as a share of their population and resulted in Israel being the only exception to "all economically developed societies have below replacement fertility." Maybe this tells us something deep, and they will one day make up the entire population. Or maybe as they spread across the globe and become their own more diverse group they will follow the same pattern as the rest of us.

[9] Indeed, we wrote a whole academic paper on this (Arenberg et al., 2022)!

include the lives of potential people, one might imagine an important cultural change towards desiring and socially rewarding higher fertility.

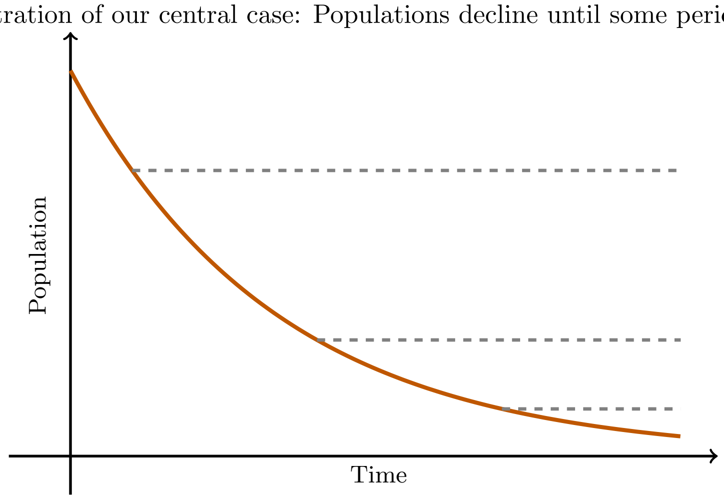
Life (+ fecundity) extension. The futurist technology that we are most optimistic could stabilize/increase populations would be increases in healthy lifespans. It remains true that—barring immortality—every individual would need to replace themselves, on average, over their lifetime. But with a 150-200 year lifespan, the opportunity cost of your marginal hours are lower. Perhaps individuals will work a full career in the first phase of their adult lives with child-raising being a standard second phase.

One hole in this story is that the most promising life-extension technologies do not appear, as far as we know, to change the speed at which a female loses her ability to bear children. This is where we believe much better reproductive technology could be a crucial complement. If something increased people's desire for children (e.g., longer lives or better parenting situations that reduce time costs), then technologies that help them achieve that desire could become instrumentally very valuable.

1.1.3 Quantitative Implications: Future populations smaller than 10 billion seem likely

To summarize the above: replacement rate fertility may be challenging to achieve. After all, every economically developed society aside from Israel has converged on low fertility. Conversely, it is also hard to imagine populations *not* stabilizing!

Figure 1: An illustration of our central case: Populations decline until some period of stabilization



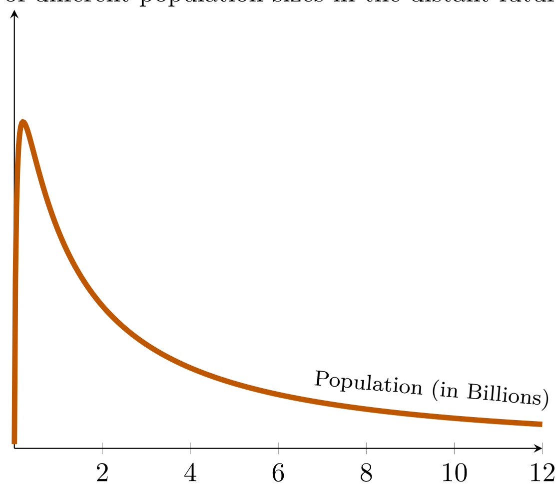
Notes: A possible path for populations given medium-run low-fertility. The zero on the x-axis represents sometime mid-22nd century, by which point population decline seems likely. This decline continues until some eventual stop. Randomness over the time/population size at which replacement rate fertility is reached provides a number of possibilities for where decline stops. Not pictured here is the possibility of rebound; though we **do** include the possibility of growth in the full space of outcomes (see Figure 2).

Assessing the chances of sustained growth is obviously difficult as well. Suppose society develops some extremely easy way to produce people, such that it takes very little social/governmental effort to prevent depopulation. Do we have any idea where such a society would grow to? For example, if a technology was developed when the world population was 2 billion, should we expect that world to get to, either by individual or social choice, 15 billion, 20, 30? Readers of this document—who we suspect are altruistic and more partial to Total Utilitarianism than the average person—might expect society to eventually share those values and aim for a world full of happy people. But it may not. For example, in a fully automated future there may be a sharp trade-off between the quality and quantity of lives. If some fixed number of AI runs all economic activity such a trade-off arises because production is invariant to the number of people. A larger population would mean forgoing quality of life, implying there would be strong incentives to prevent population growth.

All things considered then, Figure 1 illustrates the central case we consider. Populations decline until a social or technological change occurs—with some random arrival rate—that halts that dynamic. The illustration implicitly assumes we only stabilize at these different levels, but the final credence distribution (Figure 2) puts positive weight on rebound and growth (e.g., by putting weight on population sizes above 10 billion).

The final step of fully specifying probabilities over the domain of possible outcome is done as follows. First, we assume the outcomes follow a log-normal distribution. This is done to tie our hands to a well-known distribution with the following desirable properties: (1) population is bound below by zero, (2) the probability mass is concentrated non-symmetrically towards lower outcomes, (3) it allows for a long right tail of high-population outcomes and (4) it can be fully parameterized with just two parameter choices. The two parameters that we choose are the two ends of an 80% confidence interval (i.e. the 10th and 90th percentile), which we set as 500M and 50B. It is important to stress that this has not been done with a formal model, nor any explicit theory of why, for example, there is a 10% chance we end up with less than 500M people in the long-run. We do not believe current facts provide much useful evidence as to when fertility rates will get to or above replacement and, if they do recover, how many people then-current generations will to bring into existence. If one has a stronger prior about the possibility of large long-run populations, this will of course change the results of the numerical exercise.[10]

Figure 2: Probability of different population sizes in the distant future at a moment in time



Notes: Assumed credences over the number of people alive at any given future point in time. We use a log-normal distribution for transparency and because of several nice properties that accord with our views regarding the threat of persistent low fertility. Admittedly, the peak is at a lower population than we might endorse as log-normal distributions increase very quickly in mass from zero; this is (potentially) visually misleading as to where the other percentiles lie. For reference, some select percentiles of this distribution are: (10th, 500M); (33rd, 2.5B); (50th; 5B); (66th; 10B); (90th, 50B).

We conclude this section by using this distribution to illustrate an important feature mentioned throughout this report: the tail outcomes drive the mean. The median of this distribution is 5B, but the mean is 25B. In fact, roughly 80% of observations fall below the mean. So, depending on how much one wants

[10]An important simplification of the computational companion notebook is that the log-normal distribution does not allow for multi-modal distributions. As just one example, if conditional on any future population growth you expect the world go get tremendously large, that would imply more weight on 50 billion than 10 billion; something not possible in our framework. However if you put substantial uncertainty on what size a "growth-scenario" caps out at, that could be approximated by tails of a log-normal.

to chase tail outcomes, it might be reasonable to act on the statement "expected future populations exceed 20B." However, it would be equally true to state "most futures have fewer than 5B people." In the companion computational notebook we let users truncate the upper tail of final outcomes, as a quantitative implementation of anti-fanatical decision making.

1.2 Long-run rates of x-risk and expected survival

We do not have much to say about the quantitative level of existential risk we should expect to face over the coming millenia; other than, perhaps, that we should be very uncertain. If you believe (1) existential risk is high now and (2) the best guess of what the future looks like is the present, then we might think long-run existential risk remains high. Even if AI is a uniquely risky technology, Toby Ord's *Precipice* estimates ≈3% chance of extinction from "other anthropogenic threats" this century.[11] If that's the true background rate of risk and it never significantly declines, we almost certainly do not have a long, flourishing future ahead of us. On the other hand, if we're in a time-of-perils and existential risk will be persistently low after a few centuries, we might survive a very long time.

Where we differ from standard discussions of existential risks is that, because we put a higher probability on very small future populations, we put correspondingly less probability on extremely safe futures. If populations get very small it becomes much less likely that we reach a state of technological maturity, expand through the galaxy, maintain a complex modern economy, etc. Humanity may never achieve the features that we sense are important for generating persistently low existential risk if the population gets much smaller than it is today.

Beyond adding that viewpoint to the discussion, we think it is again helpful to illustrate properties of the resulting distributions. It turns out that the number of centuries we survive has an even more extreme tail that drives the large expected value. Views on fanatacism will be even more critical when assessing what to make of this distribution.

1.2.1 Two levels of uncertainty generate a heavy tail

The reason such a long tail is generated is simple: the number of centuries survived is bounded below by zero, so the more uncertainty we add, the further out the tail goes. Here there are two levels of such uncertainty that interact to produce an even more extreme tail.

1. Uncertainty over the level of existential risk.

2. The year in which we go extinct given some level of existential risk.

Let's start with a quantitative demonstration of the latter. Take a 1% existential risk per century. It's easy to analytically show that this gives us 100 expected centuries left.[12] However—something we didn't fully appreciate before running quantitative simulations[13]—was just how much this is driven by tail outcomes. Many more than 50% of futures with a 1% risk have fewer than the expected 100 centuries (64% to be exact). It is the lucky worlds where we make it 500-1000 centuries that really drive this expected value. This issue is exacerbated when we put uncertainty over the existential risk level itself; a small number of futures go on for an *extremely* long time.

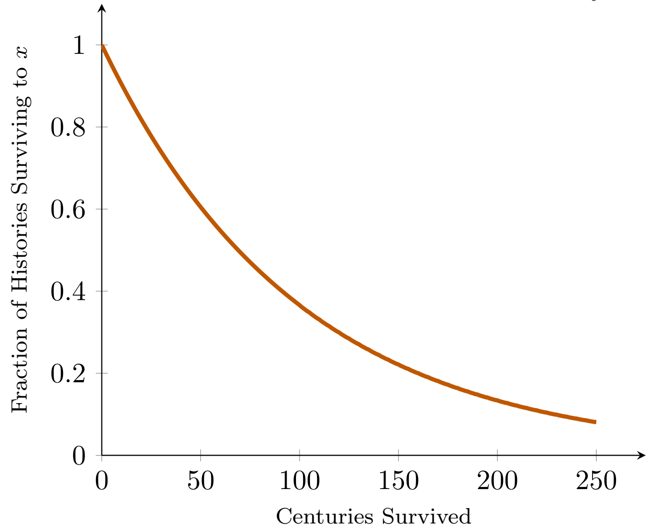
To be precise, we quantify our credence for rates of existential risk using a logit-normal distribution, again to tie our hands to a well-known distribution with few parameters. Logit-normal distributions have the nice property of looking like log-normal for small values—which is where our distribution will be centered—but are strictly bounded between zero and one. Our 80% confidence interval over this value (per-century) is [.01%-2.5%]. The lower value corresponds to an expected value of 1,000,000 years survived, commonly cited

[11] This doesn't include bioweapons, AI, or climate change; it is his category for unknown-unknowns.

[12] Or just consider the intuitively stated result: if there is a 1 in 100 chance of an event occurring in any given period—in this case, an existential catastrophe—the expected periods of survival is 100.

[13] Formally, each "history" is a draw from a geometric distribution with a given existential risk. A geometric distribution represents the first time you would draw a 1 from a 0,1 distribution with probability x on 1.

Figure 3: Illustrative Survival Curve for 1% Per Century X-Risk



Notes: This figure is meant to demonstrate the distribution of (cumulative) survival probabilities for a given existential risk. Here, just because it is a round number, we plot the survival rate for a 1% existential risk (per century). For example, 50 percent survive 67 centuries or fewer; whereas 10 percent survive more than 230 centuries.

as the length of survival for the average mammalian species. The upper bound is near Ord's estimate of unknown-unknowns this century. This distribution has a median a bit larger than 0.15% per century.

1.3 Long-run average quality of life

We treat quality of life as a second-order issue. It seems much less likely to vary by an order of magnitude than population or number of centuries survived. Recall, we are focused on scenarios that do not include digital sentience—the scenarios where very, very good lives (or very, very bad lives) seem most plausible. In a biologically based future, we might think it's possible that our internal lives do not change very much. Relative to the rapid progress in material living standards there is a case to be made that our subjective experiences have not improved nearly as much. Therefore, and partly in the spirit of conservatism, we do not assign a high probability to lives being many times better than lives now.[14,15] As a baseline, we again use a log-normal distribution, assigning 80% confidence intervals to [0.75,5.0]. In other words, we assume there is a 90% chance the life is at least as good as three-fourths of our current quality of life; and a 90% chance life is less than 5 times as good as our current lives. This generates a mean and median future where life is about twice as good as it is now, but a right tail where life ends up quite a bit better with low probability.

2 Simulations of Futures: Generating a full distribution of possibilities

Here we describe the methods used to generate the outcome of interest: a full distribution of futures conditional on the assumptions above. We show that under these assumptions the expected value of the future

[14] An exception to this is if we are currently very close to the neutral level, implying there is barely any value now. Then even modest improvements could result in a 10x-ing.

[15] Also, note that this implicitly defined as instantaneous utility, so that doubling life-spans is not a doubling in quality of life. We've folded assumptions about lifespans into our population assumptions—a doubling of lifespans increases welfare by increasing the number of people on the planet experiencing whatever instantaneous quality of life that we've assumed.

is indeed more than 10,000x the value of the present century. Because the tails drive these results, we acknowledge that the expected value does not tell the whole story. The median is only on the 100x order of magnitude; a large majority of futures fail to reach even a 1000x threshold. What one does with this information depends on how much one wants to chase tail outcomes, and how much one believes the tail assumptions of our distributions are reasonable (e.g., they don't converge to zero probability faster than a log-normal assumes, for example).

2.1 Monte Carlo simulations

The Monte Carlo simulations we run are simple to understand.

1. Provide confidence intervals and a distribution type for population size (log-normally distributed), x-risk (logit-normally distributed), and quality of life (log-normally distributed).
2. Randomly generate 1,000,000 values from each of these distributions (implicitly we are assuming these distributions are independent of one another in the implementation).
3. Use a geometric distribution to randomly generate the number of centuries it takes to go extinct using each scenario's rate of x-risk.
4. Calculate the value of the future in each scenario in terms of quality-adjusted life-years experienced. This equals 100 (years) $\times$ the number of centuries survived $\times$ average population size $\times$ average quality of life. We divide this by an approximation of the number of people alive the century starting now: 10 billion $\times$ 100 (years) $\times$ 1 (normalized quality of life) to obtain our units of "as multiples of this century".

With this numerical distribution we can report any given percentile of outcomes, compute the mean, etc.

2.1.1 Correlations across variables

One important factor we omit here is correlation across variables (see step 2). For example, our true belief is that the small population futures are the ones disproportionately likely to have high levels of existential risk and low values for quality of life. This would increase the variance, and further drive out the right tail, but should not importantly impact the median of the distribution. We'd be making the good scenarios even better, and the bad scenarios even worse. Because this would not change the qualitative takeaway—that the expected value would remain over at least 10,000x the present century, but the median value would be between 100-1000x—we do not complicate the model in this way.

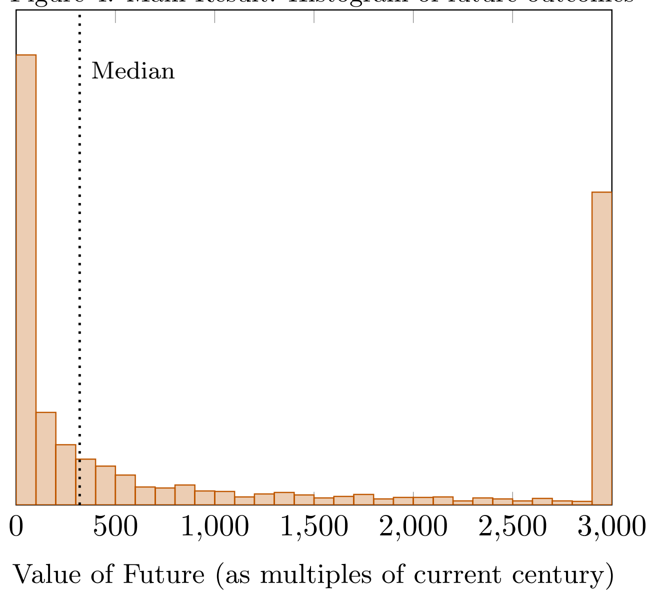
2.2 Means, medians and fanatical assumptions

The final results are plotted in Figure 4. As we previewed throughout the document, the distribution over possible future values has a long right tail. Indeed, the fraction of histories with more than 3000x the value of the present century is nearly as large as the very low-value scenarios (that tall bar on the right-hand side truncates all values above 3000x for visual purposes). This broad pattern of a long right tail dominating the expected value seems like it will be arise for most distributions with sufficient uncertainty because the key inputs have non-symmetric distributions.[16] Specifically, with the assumptions we fed in, the *expected* value of the future is about 40,000 times the value of the present century. The median value of the future is only 320 times the value of the present under these same assumptions.

What one chooses to do with this information depends on one's decision-theory, and trust in the tail outcomes. If we are reasoning in orders of magnitudes around central cases, then the fact that nearly 70% of futures generate less than 1000x the value of this century may be the relevant empirical claim when considering whether it is "conservative" to treat the value of the future as containing at least 1000x the

[16] A possible exception would be if the distribution of average quality of life was symmetric around zero; e.g., if the possibility of very bad digital futures was about equal to the possibility of good futures.

Figure 4: Main Result: Histogram of future outcomes



Notes: This figure plots the distribution of outcomes for future values (computed as described in Section 2.1). Values are reported as quality-adjusted life-years relative to the present century, so that a value of 500 implies those simulations hold 500x the quality-adjusted life-years of this century (approximated as 10 billion x 100 years). The x-axis is restricted to 3000x because otherwise the tails swamp the image; all of outcomes above 3000x are bunched in the image (but not expected value computation).

value of the present century. Conversely, the large expected value seems quite robust to a wide range of parameters. That is a fairly compelling reason to believe the future does contain a lot of *ex-ante* value.

One tool that we have included in the spreadsheet is an "anti-fanatical" slider. This lets the user truncate the distribution at the top p percentile. For example, if you choose $p = 1\%$ everything in the top 1% of values gets manually reset to the 99th percentile value. This reduces the quantitative influence of these observations, without throwing them away. Because the tails have so much of the value, this only ends up bringing the expected value of the future below 1000x if one truncates a large fraction of futures (10%+). This could become an important consideration for different distributional assumptions.

3 Conclusion

The expected value of the future appears to be very large even when we set aside the possibility of digital consciousness and truncate the tail outcomes to reduce the influence of a small number of extremely good outcomes. However, we believe there is a reasonable probability mass on low-value futures even if we do survive near-term existential threats. Low-fertility threatens large, flourishing societies; randomness, even in the case of relatively low long-run x-risk, threatens to cut off our history early, even if the expected number of centuries we ought to live is large.

More research on the distributions of these variables would of course refine these numerical solutions. For example, it was beyond the scope of this attempt to search the space of distributions for one that has a shape more faithful to our true credences. While log-normal and logit distributions have some useful properties, they have some irregular features that could potentially be eliminated with a more complicated distribution (Weibull, etc). However, we feel fairly confident that the qualitative takeaways of this report will be fairly robust: the expected value of the future will be large absent strong confidence in high levels of existential risks, but this expected value may mask that many, or even most, futures contain comparably little value.