What We Owe The Future

Introduction

As I write these words, the United States is witnessing the largest upswell of political protests in its 250-year history. Led by the Black Lives Matter movement in response to the killing of George Floyd by policeman Derek Chauvin, these protests have brought renewed attention to racial injustice in the United States and elsewhere.

The debate that has broken out in response to these protests has been predictable: progressives demand better treatment of Black Americans; conservatives ask whether it's really so bad to be Black in modern America.

What is under-discussed is just how deep the roots of the problems faced by Black Americans go. When we talk about the problems of overincarceration and police brutality against Black people, what we should be talking about is not merely the issue's proximate causes like poverty, the breakdown of the family unit, or racism. We should be talking about slavery.

Slavery was 'abolished' in the United States at the end of the civil war in the mid 19th century after the victory of the Northern states. This caused an economic problem for the Southern states, which had been heavily reliant on slave labor: the 4 million slaves that were freed had accounted for 50% of their wealth and 34% of their population. How would Southern landowners be able to operate their vast cotton plantations in the absence of slaves?

They found their solution in the 13th amendment, which made slavery unconstitutional but contained a crucial loophole: allowing slavery 'as a punishment for crime.' So the white elites did as any profit-maximizing and racially prejudiced set of actors would: they implemented a set of laws, now known as the 'Black Codes,' that enabled imprisonment for minor offences, criminalized vagrancy (that is, having no evidence of employment), excluded Black people from juries, and set up a system of convict leasing, where state prisons could hire out convicts to work on plantations. The loophole allowed the slavery of Black people to continue.

This practice was trenchant. Up until the 1870s, the Mississippi State Penitentiary modelled itself on an antebellum cotton plantation: they purchased 20,000 acres of land, racially segregated the inmates, and set them to work farming or picking cotton, often forcing them to work 15 hour days in intense heat. They implemented a 'trusty' system, where lifelong inmates, therefore typically the most violent criminals, would supervise the other inmates who were picking cotton; the trusties had the power to dole out corporal punishment, including beatings with a leather whip known as Black Annie, and the most senior trusties were given guns. The penitentiary was highly profitable, making $5 million in today's money for the state in its first two years of operation. In those same two years, 7% of the inmates died.[1]

These horrors might seem distant to us now; the past is a foreign country. But I mistyped a number in the previous paragraph. It wasn't the 1870s when the Mississippi State Penitentiary stopped its practices; it was the 1970s. De facto slavery was practiced in parts of

[1] 'Worse Than Slavery', Oshinsky

the US as recently as 50 years ago. And even today, there are over 500,000 prisoners in the United States engaged in penal labour — all because of an idiosyncratic decision made by the authors of the 13th amendment in 1865. Today, the USA's incarceration rate is five times that of comparable countries like the United Kingdom, and Black people make up 40% of the prison population, even though they constitute only 6.5% of the US as a whole.[2]

The narrative I've given you isn't just stylised storytelling, but has recently been studied quantitatively. Economist Melissa Rubio studied US census data from 1850 to 1940: Black incarceration rates increased substantially immediately after the abolition of slavery, and continued to grow thereafter. The greater the proportion of a Southern county's population that had been enslaved, or the better the suitability of its land for growing cotton, the higher the incarceration rate of Black people in that county after abolition; this remained true and statistically detectable up until 1940, until confounders like the Great Migration of six million Black Americans out of the South into other parts of the US caused too much noise in the data. What's more, she found that in counties that had lower cotton output as a result of the 1890s boll weevil plague, Black incarceration rates were lower, suggesting that high incarceration rates were at least in significant part driven by the demand from white elites for cheap prison labor.[3]

And it's not just incarceration rates that have their roots, in part, in slave-owning practices centuries ago. A variety of econometric studies have looked at other aspects of the long-run legacy of slavery. They have found evidence that Southern counties with higher rates of slave ownership before abolition have higher rates of economic inequality between races today,[4] more violent crime,[5] more hate crime,[6] higher rates of racist attitudes,[7] more race-related constitutional violations,[8] lower rates of educational attainment among Black people,[9] lower rates of support for affirmative action,[10] and lower Black voter turnout.[11] Counties that had more frequent lynchings in the late 19th and early 20th centuries have lower rates of voter registration among Black people today,[12] and higher rates of death at the hands of police officers.[13]

To understand the problem of racial injustice in America, we would be missing the mark if we just looked at the current obstacles that Black people face. We are born into a society

---

[2] Rubio, https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=ESWM2019&paper_id=666
[3] Rubio, https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=ESWM2019&paper_id=666
[4] https://izajodm.springeropen.com/articles/10.1186/s40176-016-0072-0,
[5] https://papers.sioe.org/paper/1700.html
[6] https://www.researchgate.net/publication/338680277_The_Legacy_of_Slavery_on_Hate_Crime_in_the_United_States
[7] The Political Legacy of American Slavery, Avidit Acharya, Matthew Blackwell and Maya Sen
[8] A Culture of Disenfranchisement: How American Slavery Continues to Affect Voting Behavior, Avidit Acharya, Matthew Blackwell, and Maya Sen
[9] http://ftp.iza.org/dp5329.pdf
[10] The Political Legacy of American Slavery, Avidit Acharya, Matthew Blackwell and Maya Sen
[11] A Culture of Disenfranchisement: How American Slavery Continues to Affect Voting Behavior, Avidit Acharya, Matthew Blackwell, and Maya Sen
[12] https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxqaGF3aWxsaWZpbGVzfGd4OjJmMGMyMDY2NjcwMzQyMTI
[13] https://inthesetimes.com/article/22618/police-killings-black-lives-matter-lynchings-george-floyd

structured in part by the atrocities of our forebears, and to understand the predicament of Black Americans today, we need to look back centuries.

What's more, it's not just the United States that bears the scars of slavery. Over 20 million people were taken from Africa in the Atlantic and Arab slave trades, and suffered almost unimaginable misery. But the startling fact is that the greatest harms of the slave trades came after its abolition. A significant part of the reason why countries in sub-Saharan Africa today are so much poorer than even other developing countries is because of the civil wars, ethnic fractionalisation, and breakdown of social trust that resulted from the slave trades.[14] Though at the time the slave trades killed millions, and caused tens of millions of lifetimes of misery, the long-run effects of the slave trades will consign billions to starvation, poverty or early death.

Crucially, it was not inevitable that slavery would be abolished. As I'll argue in chapter 4, it's quite plausible that, had political circumstances been different in the 18th and 19th centuries, or had the small group of Quaker activists that founded the abolitionist movement never existed, there would still be widespread legally-sanctioned slavery today. The example of the abolitionists shows us that, at least in some cases, it is possible for our actions to have a lasting positive impact on the future — potentially improving the world in ways that will reverberate for centuries to come.

The persistent problems of inequality between Black and white people in America — symbolized in the image of a white police officer's knee on the neck of a passed-out Black man, and the words "I can't breathe," — give us an all-too-visceral illustration of the past's grip over the present. The future is long, and the greatest harms of moral crimes and ignorant mistakes often lie in the shadow they cast over the centuries that follow.

And it's not just slavery whose legacy we live under; our world is shaped in innumerable ways, both negative and positive, by actions taken in the distant past. This is true of the natural environment: North America is known for its plains and savannas, but it was covered in forests until the first humans migrated there twenty thousand years ago and burned them down. We associate megafauna like giraffes, elephants, and rhinoceroses primarily with Africa, but tens of thousands of years ago the entire world had similar animals; South America alone hosted four-tonne ground sloths, 70-kilogram dire wolves, and a relative of the armadillo the size of a car. It was because of human hunting that these species were irrevocably killed off.

The distant past has also shaped present-day technology. I would not be sitting here typing on this laptop if not for the thousands of advances made cumulatively over millennia: computing and electricity of course, but also plastics, batteries, manufacturing, all the way back to metallurgy, glass-making and writing. Much the same could be said of most of my daily activities. It's certainly true of the foods I eat: the bulbous and fleshy fruits I buy at the supermarket without a second's thought are the product of centuries of selective breeding. The humble potato was domesticated in the Andes 10,000 years ago and imported to Europe

[14] Nunn, The Long-term Effects of Africa's Slave Trades; Nunn and Wantchekon, The Slave Trade and the Origins of Mistrust in Africa. [not sure if I can stand by this causal claim.]

and Asia in the 16th Century; the greater agricultural yields it enabled were so revolutionary that without it, on one estimate, a billion fewer people would be alive today.[15]

Similarly, political institutions can have long-lasting legacies. The Indian caste system dates back more than two thousand years, and is still responsible for segregation and discrimination to this day. Exploitative institutions set up in the colonial era have persisted, producing poorer modern nations.[16] The fact that I am lucky to live in a parliamentary democracy is due, ultimately, to the British Revolution of 1688.

Finally, the present day is often shaped by the culture and values of the distant past. 60 billion animals are killed for food every year, most of them in factory farms, in part because it was Christian morality that became dominant rather than vegetarian-friendly Hinduism. Similarly, we live in a world currently dominated by individualistic European culture and moral norms — a culture that is an outlier on many dimensions — because that's where industrialisation first occurred, 250 years ago. One recent study argues that this individualism can be traced back all the way to the Catholic Church prohibiting marriage between cousins in the Middle Ages.[17]

Previous generations have made today worse in many ways, and they have also bestowed on us enormous gifts. Presented with the wide variety of ways in which the actions of the past have positively and negatively impacted the world today, we should ask: what are we in the present generation doing that will have long-lasting effects into the future? And what should we do differently if we want to ensure a better world for those generations that are yet to come?

That's what this book is about.

This book is about a moral viewpoint that I and others have developed over the last two decades,[18] which I call *longtermism*: a view that is particularly concerned with future generations, and which wants to make things better not just in the present but also for the centuries and millennia to come.

It took me a long time to come around to the longtermist perspective; it's hard for an abstract ideal, involving concern for generations of people who do not yet exist and whom we will never meet, to have the same motivational force as the problems we can see in front of our eyes. As a teenager, I was motivated by the hardship around me: I ran summer camps for low-income disabled children in Glasgow; I worked as a carer for those suffering from dementia in an old folks' home. Over time, the arguments for taking a global, cosmopolitan perspective won me over, and, appreciating the overwhelming scale of global inequality, I resolved to figure out how individuals could best help those in extreme poverty. I committed to give away at least 10% of my income to charity, and I cofounded an organisation, *Giving What We Can*, encouraging others to do the same. After reflection I subsequently aimed to

---

[15] Nunn and Qian, The Potato's Contribution to Population and Urbanization, plus extrapolation from 1900 to the present day

[16] Acemoglu, Johnson, Robinson, The Colonial origins of comparative development; Acemoglu and Johnson, Reversal of Fortune; etc.

[17] Schulz, Jonathan, et al. "The origins of WEIRD psychology." Available at SSRN 3201031 (2018).

[18] [putting publication of Astronomical Waste as the start date. Could go further back to Reasons and Persons, or earlier?]

give away everything I ever earn over approximately the typical full-time income in the United Kingdom, which is about £30,000 per year in 2020.

It was in the course of engaging with others who shared my goal of helping people as much as possible, in what we came to call the effective altruism community, that I was first presented with the arguments for taking the perspective of the very long run, and for thinking that the flourishing of future generations should be a, if not *the*, moral priority of our time. My instinctive reaction was one of confusion and dismissal. Why should I care about what happens in hundreds or thousands of years hence? And, even if I did care, what on earth could we do about it anyway? But the strength of the arguments in favour of longtermism had an inexorable pull. Some simple ideas drove this: that future generations should count for no less, morally, than the present generation; that the human story might be very long indeed, and so there is a vast amount of potential yet to come; that the future could be extraordinarily good or inordinately bad, and that it may be within our power as a society to choose which; and, most of all, that future generations are utterly disenfranchised in the world today, lacking a vote or lobbying power or any other means to represent themselves, so it is left up to us to give them a voice.

I came to see concern for future generations as continuous with the ideals of the social justice movements of the past and present that have inspired me, like abolitionism, civil rights, women's rights, and cosmopolitanism. The history of moral progress is, in part, a history of expanding one's circle of concern; progressively coming to give due moral consideration not just to one's friends and family, but also to strangers, and people of other races, religions, and nationalities. Taking the interests of future generations seriously seemed to be the next stage of moral progress.

I have tried to write the book that would have convinced me, a decade ago, of longtermism. The core claims I make in the book are these. First: future people count, morally (chapter 1). Moreover, in expectation, they are vast in number, yet they are utterly disenfranchised. This means that *if* we can predictably impact the long run, then doing so is one of the most important things we can do. Second: we can indeed predictably impact the long-term future (chapter 2). One effective way to do so is via promoting clean energy innovation as a means to combat climate change (chapter 3). This activity is unusual in how robustly good it is from a long-term perspective. However, other actions, though based on more speculative arguments, are more neglected and therefore have the potential to be even higher-impact. These include social activism to improve society's values (chapter 4), shaping the development of advanced artificial intelligence (chapter 5), and preventing risks to the survival and flourishing of civilisation, in particular as a result of war between great powers (chapter 6).

In this book I focus on making the case for longtermism, and presenting the range of options available to the longtermist, rather than arguing for one particular priority. Indeed, there are different species of longtermism, depending on which big-picture judgment calls you make, which can affect which activities you regard as most important.

First, you can be *bold* or *cautious*. If you are cautious, then you will be comparatively more inclined to take actions that are supported by a large body of evidence, and that are beneficial

across a wide variety of ways the future could go.[19] In contrast, if you are bold, you are more willing to act on more speculative arguments that promise an even larger impact.

Second, you can be *procreationary* or *noncreationary*. Do you think that, if civilisation is good, it is better for that civilisation to be larger or longer-lasting, thereby involving more people with flourishing lives, more knowledge and art and accomplishment? Or do you think the size of a civilisation is a neutral matter — that we should only prefer to make civilisation better at those times and locations at which it will exist either way? Procreationists endorse the former; noncreationists the latter. I discuss this issue in chapter 7 on population ethics.

Third, you can be *optimistic* or *pessimistic*, representing how good you think the future will be if civilisation continues to grow for a long time. If you think it will be very good, then you will be more inclined to ensure that it happens at all, and more likely to prioritise reducing the risk of human extinction, permanent civilisational collapse, or technological stagnation. If you think it will be only ok — good on balance, but far short of how good it could be — or even bad, then you will be more inclined towards trajectory change, trying to improve the value of civilisation in the scenarios in which civilisation lasts a long time. I discuss the issue of the value of the future in chapter 8.

Fourth, you can be *patient* or *impatient*, depending on whether you think that we can best improve the long-term future by taking action today, or instead think that we will have more influence at a later time and should invest our resources until then. I discuss this issue in chapter 9.

Personally, I think that the arguments favour *moderate boldness* (where we should be willing to act on more speculative arguments, but we should do so only reluctantly)*, procreationism*, *moderate optimism* (expecting the future to be good on balance, but very far from its ideal state) and *patience*. But none of these are settled issues, and I will try to present an even-handed discussion, though never presenting arguments I regard as weak.

I appreciate the enormity of the challenge of not only defending the view that we should be seriously thinking about the impact of our actions over centuries, millennia, or even longer, but also giving a comprehensive account of what's most important from that perspective, trying to get the big picture at least roughly right. I didn't take this challenge lightly, and my solution was to rely heavily on an extensive team of consultants and research assistants, who are themselves consolidating and building on decades of cutting-edge research from many different academic fields, and on research done within the effective altruism community. By necessity, I have had to cover a broad range of disciplines: not only my own speciality of moral philosophy, but also history, political science, economics, forecasting, cultural evolution, climate science, international relations, and artificial intelligence. In every case where I have stepped outside my subject-matter expertise, I have relied on experts in the field from the early stages, and ensured that the relevant sections of the book have been reviewed by at least two domain experts.

---

[19] In this book, I don't discuss whether we should be bold or cautious, but articles that do discuss this issue include: [list, including GiveWell, RAND stuff, something on broad vs narrow interventions, and Mogensen / Thorstand's work]]

This book is therefore not really ‘my’ book: it is the product of many lifetimes’ worth of work, spread over dozens of people. For those who want to dig deeper into some of my claims, I have compiled a list of supplementary materials, available on the book’s website: often these are articles by myself or colleagues, or technical reports that I have commissioned in the writing of this book. Together these supplementary materials amount to over half a million additional words.[20]

Despite these efforts, I believe I have only scratched the surface of what could be said about longtermism and its implications. If the sole consequence of this book is that it serves as a foil for others to explain why its claims are misguided, I will consider it a success.

[20] [This is my guess at how long it will end up being. Could well be more.]

Chapter 1: The Case for Longtermism

§1 The totality of human experience[21]

Imagine if you had the opportunity to experience, in order, the lives of all 100 billion people who lived in the past and all 8 billion people alive today. As you would experience it, your life would start 200,000 years ago, in eastern Africa, with the birth of the first ever human being; when that first human died you would be reincarnated and wake up as a newborn, inhabiting the body of the second human ever born, then the third, then the fourth, until the death of the most recently born person today.

Over your many lifetimes, you would experience a very slow and faltering accumulation of technology: the gradual improvement of stone axes, spears, clothing and pigments; then the inventions of farming, pottery, domestication of animals, smelting of metals, and textile-making. You would witness the squeezing out of hunter-gatherers by their agricultural brethren, and then the conquest of agricultural societies by gun-toting proto-industrialised colonisers. You would see empires rise and fall; statues of God-Kings turned to dust.

You would feel immense pain and, occasionally and increasingly, intense joy. Of the 1.65 trillion life-years you would live through, you would spend 1 billion years giving birth, 7.5 billion years having sex, and 50 billion years drinking coffee.

Because of increasing population sizes, half your life would come after 1300AD, and a third would come after the industrial revolution, when technology and societal organisation start to change far faster than they had ever done before. You would witness the development of steam engines, factories and electricity; multiple revolutions in science; the most deadly wars in history; extensive destruction of the natural environment; the harnessing of animals for food on an unprecedented scale; and rising living standards, with hundreds of millions of people now living like the kings of yesteryear. 15% of your life would be the experience of people alive today.

That would have been your life from the birth of *Homo sapiens* until the present moment. Now imagine that, in addition to all past and present lives, you will also live the lives of all those who will exist in the future. By the time you reached today you would have already experienced over a trillion years, but your life would be just beginning. Though we cannot know for how long humanity will persist, we have good reason to think that the future will be vast. Even if world population decreased to a tenth of its current size, and even if humans lived merely as long as the typical mammalian species, then 99.8% of your life would still be left to come; if such a lifetime of humanity were compared to a typical human life, you in the present day would be just 7 weeks old. If humanity survived longer than this — the hundreds of millions of years until the earth is no longer habitable, or the hundreds of billions of years until the last stars burn out — your 1.65 trillion years of experience so far would be like the first blinking seconds out the womb. The future is big.

---

[21] [Sources: Eukaryote on LessWrong, and RM Hare on ‘impartial observer’]

Now suppose that you were not merely experiencing all lives, but that you were also aware of what would happen to you: you knew that every time you died you would be reincarnated in the next human's life, and you had the ability to control and make decisions in those lives, too. If you knew that you still had trillions of years of life to live, how would you choose to act, today? Would you pollute the climate with emissions of greenhouse gases? Would you drive species to extinction via the destruction of the rainforests? How much would you invest in research and education? How careful would you be with the development of new technologies that could potentially curtail your future?

Most importantly of all, how much time would you spend thinking about the ways in which the actions of today will impact the long term? We encourage teenagers and young adults to think hard about the actions that will have consequences for the long-run course of their lives: whether to go to university, what subject to study if they do, what career to pursue, where to live and who, if anyone, to settle down with. And we encourage them to be prudent: to avoid drugs that risk addiction; to avoid dangerous activities that risk death or a criminal record; to invest in education and then, later, a pension. We do all this because forethought and prudence early on pay dividends many times over through the course of a life.

The same reasoning applies, but with ten thousandfold force, to decisions about how we in the present should think and act with respect to those who are yet to come. We are currently like a teenager, making decisions that might dramatically impact our long-term flourishing, but with a lifespan measured in the millions of years. Yet we do not act with prudence or forethought; we're short-sighted, focused on immediate gratification, with only the haziest concern for our future selves.

I give you this thought experiment because morality, in significant part, is about stepping outside of our own parochial concerns, placing ourselves in others' shoes, and taking others' interests seriously in the same way we take our own interests seriously. Though we cannot feel what it's like to live all those many lives other than our own, morality's imperative tells us to consider how we would act if we could. If we could, we would feel intense concern for those in the present who are disempowered and experience pain. But, most of all, we would be struck by the impact our actions today can have on the vast future ahead of us, where almost all potential for joy and suffering lies.

It's this impartial perspective, combined with the appreciation of the sheer scale of the future, and our ability to influence it, that leads to longtermism. In this chapter I'll make the moral case for longtermism in more depth. My primary argument runs as follows, based on four premises. First, future generations count, morally: someone's mere location in time is as irrelevant, morally speaking, as their geographical location, their race, or their gender. Second: in expectation, the future of civilisation is vast, both in duration and in the sheer number of people who are yet to come. From these first two premises we can draw the conclusion that the wellbeing of future generations is of enormous importance.

The third premise is that future people are utterly disenfranchised in the world today, lacking the vote, the ability to bargain with us, or the ability to represent their interests via activism or political protest. From this it follows that, if there are actions that society can take to benefit future generations, there is little reason to expect those actions to have been taken. The final premise is that there are indeed ways society can in expectation meaningfully affect the

wellbeing of future generations over the long run. So we can conclude that there are things we can do today that have enormous moral importance because of their impact on the long-run future. In other words: trying to ensure that the long term goes well is a key moral priority of our time.

I hope that each of these premises will seem plausible given a moment's reflection. But they are so central to the longtermist worldview that it is worth spending some time defending them in turn.

§2 Why future generations count

Some people might object to longtermism on the grounds that future people matter less than those alive today, simply because of their location in time. However, the attempt to find some characteristic by which to demarcate classes of individuals, and assign those classes greater or lesser moral worth, does not have a good track record. There is a long tradition of thinkers claiming that there are intrinsic differences between races or sexes or religious groups, and that such intrinsic differences warrant treating people outside of one's own group as morally inferior. We now regard such views as abhorrent. We know that race, sex and religious affiliation are morally irrelevant characteristics of an individual. But if these demarcations have fallen under the steamroller of moral progress, we should expect that location in time will, too. Just as racism is the idea that some races are of lesser moral worth, and sexism is the idea that one sex or gender is of lesser moral worth, we can coin the term *presentism* to refer to the prejudice that people alive at other times are of lesser moral worth than people alive today.[22]

Of course, merely coining a term does not show that this is a true prejudice. But this idea is well-supported by argument. Imagine you are hiking on a rarely-used mountain trail, and you drop a glass bottle, which smashes on the ground. You are debating whether or not to tidy up the shards: you aren't sure whether you should because people rarely use this trail, so you think the risk of someone cutting themselves on the glass is low, and it would be inconvenient to have to carry the remnants with you on the rest of your hike.[23] You're about to continue your journey, leaving the shattered glass behind, when your phone rings. The call turns out to be from someone who lives one hundred years in the future. They have just been walking on this trail, stopped for a picnic, and cut their hand on the glass you dropped. So they are asking, politely, if you can tidy up after yourself.[24]

If you got this call, what would you say to the caller? It would seem bizarre for you to respond by saying: 'Sorry, I'm still not going to tidy up. Yes, you've cut your hand and I've harmed you. But you're just a future person: you're so far away in time that I don't care about you.' Intuitively, it doesn't matter whether this phone call comes from one hundred years' time, or one thousand years' time, or tomorrow. No matter *when* the call comes from,

---

[22] [NB this framing and sentence is an homage to Animal Liberation, where Singer makes a parallel argument regarding speciesism.]

[23] [Reference Parfit for this example]

[24] [caveat non-identity issues, which I'll return to later. Maybe discuss time travel? This story is consistent with physics as we know it, though we know that the recipient of the call will not, in fact, tidy up after themselves.]

the core moral facts are the same: your action could harm someone, and, insofar as it is easy to take action to prevent that harm, you should do so. Location in time is irrelevant.

That story made an appeal to our moral intuition, to show how the view that future people matter is not morally controversial. But this view is also justified on more theoretical grounds. The thought experiment I gave you of civilisation as a single life was one way of showing that an impartial moral stance should lead us to give great weight to future generations. But multiple paths lead to the same conclusion. Imagine, for a moment, that there is a God who allots disembodied souls to bodies, and that you are such a disembodied soul, waiting patiently in the queue to be born. You do not know what body you will enter into: whether you will be white or Black or brown; whether you will be male or female or intersex; whether you will be rich or poor. Nor do you know when you will be born: for all you know, you could become a hunter-gatherer in 100,000BC, or an agriculturalist in the Roman Empire, or a space-faring adventurer born in 10,000AD. For every soul, God assigns them a body randomly chosen across time and space.

Most souls simply have to accept the lot they are given, but for you God makes an exception. It's still the case that the body you enter will be randomly chosen across time and space. But God gives you one special power. In advance of your birth, you can make decisions about how society is structured: for all different times and for all different countries, you can choose what political institutions operate, and what moral norms those alive at the time follow, in order to ensure that, when you are born, you have the best life you can have.[25]

If you were given such power, what decisions would you make about the institutions and moral norms of 5,000 BC, 1AD, or the early 21st century? There is plenty of room for reasonable disagreement on the details, but one thing is clear. Because you wouldn't know *when* you are going to be born, you would want to ensure that society is as flourishing as possible across all time, giving no preferential weight to any particular period except insofar as more people live at some times rather than others. Let's look at the 21st century as an example — what would you want society to look like today? You might note that it's much more likely that you would be born after 2020 AD, given the sheer number of people that will likely be born after that point, than to be born before or during this century. So, given the great likelihood that you would be born much further along in time than today, you would have very strong reasons to ensure that society in the 21st century is *future-oriented*: deeply concerned with ensuring that, wherever possible, society makes choices to ensure that the long-run future goes as well as it can.

Like the idea of civilisation as a single life, this thought experiment — known as the 'veil of ignorance'[26] — helps us look at the world from an impartial, third-person perspective. From this perspective, we see that one's location in time is a morally arbitrary characteristic of one's life. So location in time should not be used as grounds for discriminating between one generation and another.[27]

---

[25] [Say we can assume for now that God doesn't let you choose *how many* people are alive across time. Say we'll come back to this thought experiment, with variable populations, in chapter 7.]

[26] First proposed by the Nobel-winning economists William Vickrey and John Harsanyi, and later popularised by political philosopher John Rawls.

[27] There are many other arguments one can give for being impartial across time, and giving the wellbeing of those in future generations the same moral weight as the wellbeing of those today. [ref somewhere that makes the case, maybe HG's discounting paper]. For example, one surprising path to the conclusion that location in

§3 The size of the future

We have seen the case for taking the interests of future generations seriously. This idea, however, gets its bite when combined with the empirical fact that, in expectation, there is a prodigious number of people who are yet to come.

To see this, let's zoom out and consider the long-run history of the human race. There have been members of the genus *Homo* on Earth for over 2 million years. Our species, *Homo sapiens*, evolved around 200,000 years ago. Behavioural modernity — the use of abstract thinking, long-term planning, and symbolic behaviour such as burials and artistry — developed 50,000 years ago. We first implemented agriculture just 10,000 years ago, the first cities formed just 3,000 years ago, we entered the industrial era a mere 200 years ago, and all the changes that have happened since then — progressing from horses and carts to space travel, leeches to heart transplants, abacuses to supercomputers — occurred during what is a blink of an eye on evolutionary timescales.

---

time is of no moral significance is based on our best understanding of physics, and in particular on Einstein's theory of general relativity. [this argument is also made by Tyler Cowen] The key aspect of the theory, for our purposes, is that, fundamentally, time is no different from the three spatial dimensions. And, just as different people can move through space at different speeds, different people can move through time at different rates.

There are two ways in which this can happen: if someone else is travelling faster than you, or if they are under much stronger gravitational force than you, then they will travel through time more slowly than you will. If your sibling were to leave Earth and travel close to the speed of light, or were to move close to a black hole, and then returned what from your perspective was forty years' later, they would have barely aged at all. Indeed, by travelling quickly, or by moving close to a major source of gravity, they could return to Earth in thousands of years' time.

But this shows that location in time cannot be of moral importance. Suppose you have a young friend who has a genetic predisposition towards Alzheimer's, and there is a treatment you could pay for, now, that would prevent their deterioration in their old age. Would you regard that treatment as less valuable if you found out your friend were to travel into space at high speeds or close to a strong source of gravity? If we have the view that one's location in time matters morally, then we would have to conclude that finding out about your friend's space travels ought to change how you value that treatment. But this is absurd. What matters is the benefit you can provide to this person: just as their location in space, or how fast they are moving, is irrelevant, their location in time, and how fast they are moving through time, is similarly irrelevant.

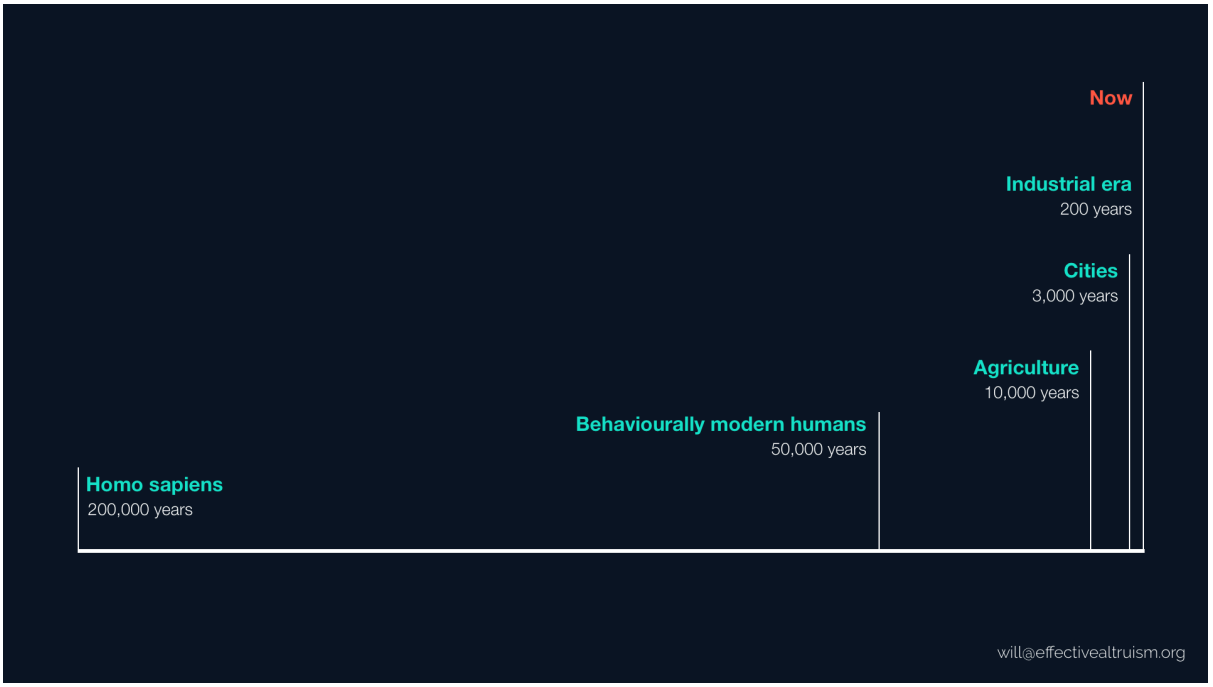


Fig XX. The history of Homo sapiens

Of course, we don't know how many years are still to come, so we have to make estimates as best we can, and take into account our great uncertainty. But, for the purposes of my argument, even a conservative estimate is sufficient.

Suppose, for example, that we last merely as long as the typical mammal species — that is, around 1 million years, and that our population continues at a tenth the size it is today. If so, there would be 8 trillion people to come, and future people would outnumber us by a thousand to one.

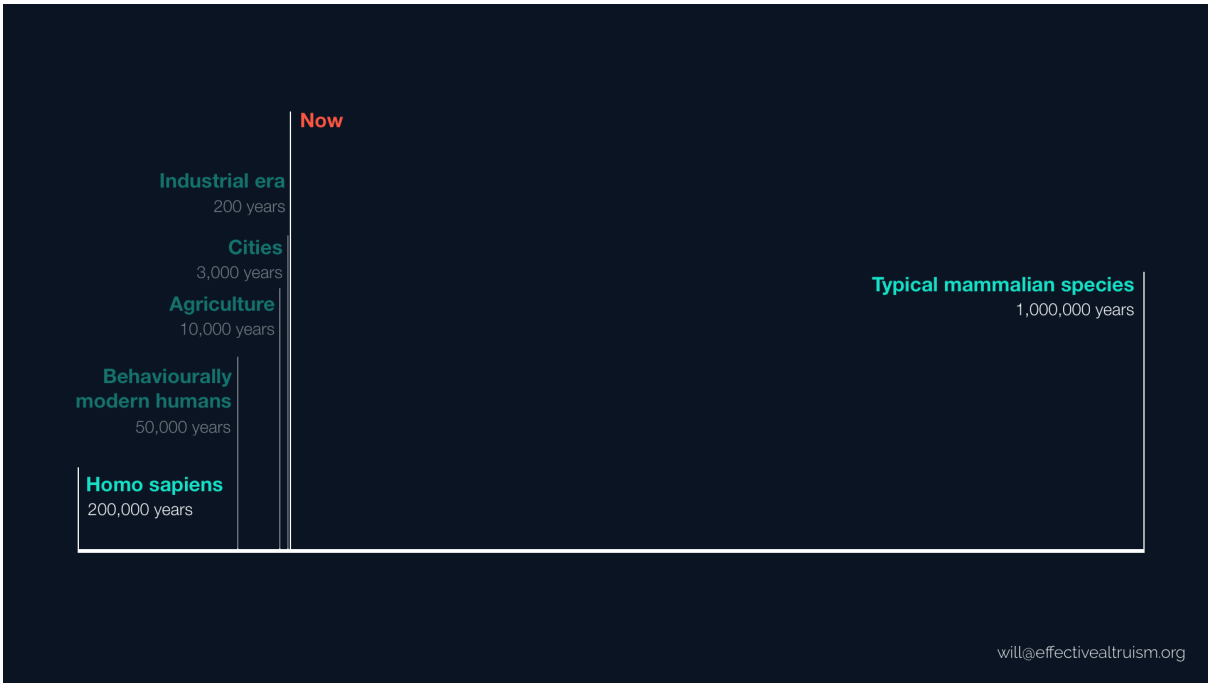


Fig XX. The potential future of civilization

But this estimate does not come close to an upper bound on civilisation's lifespan, because we are not a typical mammalian species: while we face additional challenges like anthropogenic risks of global catastrophe, we also have advanced tools and technology that give us extraordinary environmental adaptability; abstract reasoning that allows us to undertake long-term planning in response to novel circumstances; and a shared culture that allows us to function in groups of millions. Together that has resulted in a species with one hundred times the biomass of any wild vertebrate animal that has walked the Earth, across almost every environmental niche on the planet.

Given humans' abilities, there's no reason why we couldn't live much longer than the typical mammal. The Earth will remain habitable for hundreds of millions of years to come — if we were to survive that long, there would be a million people in the future for every person alive today. And if humanity ultimately takes to the stars, the timescales become literally astronomical. The Sun will keep burning for five billion years; the last conventional star formations will occur in over a trillion years' time; and, due to a small but steady stream of collisions of brown dwarfs, a few stars will still be shining in a million trillion years' time.[28]

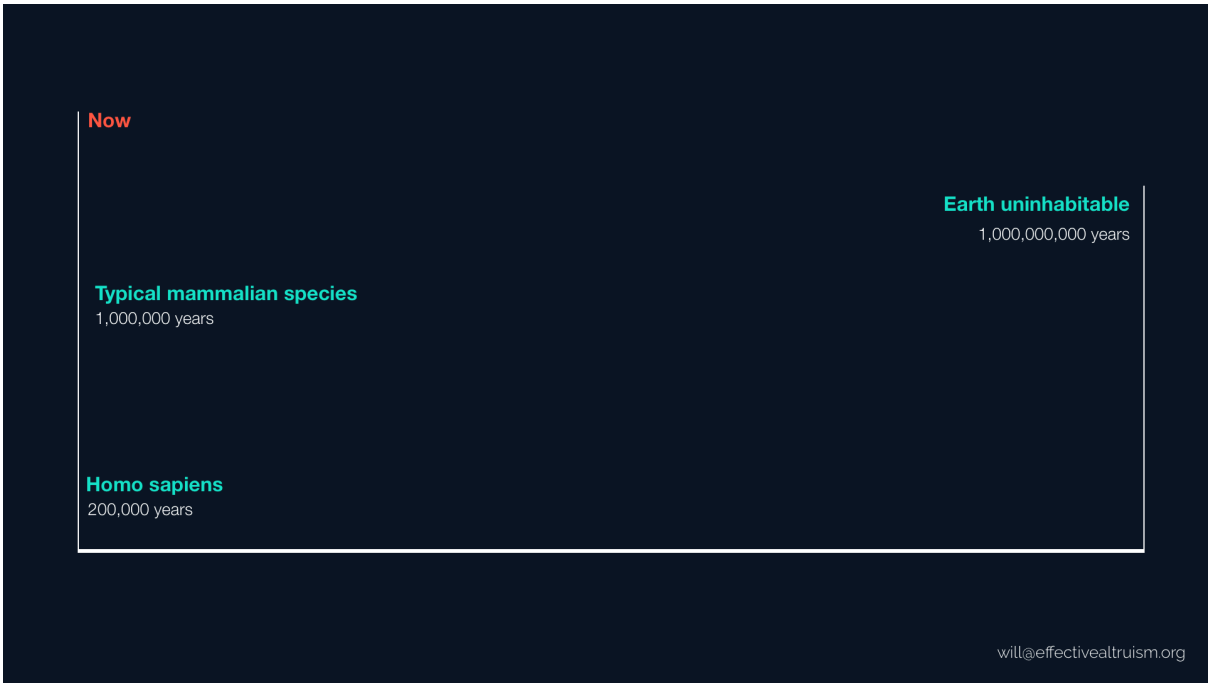
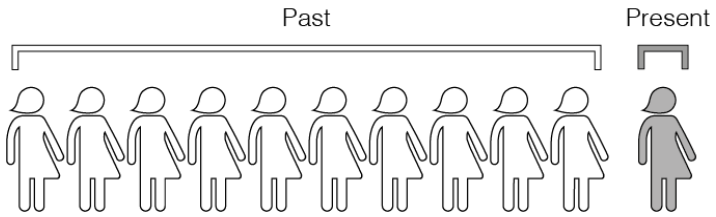


Fig XX. The potential future of civilization

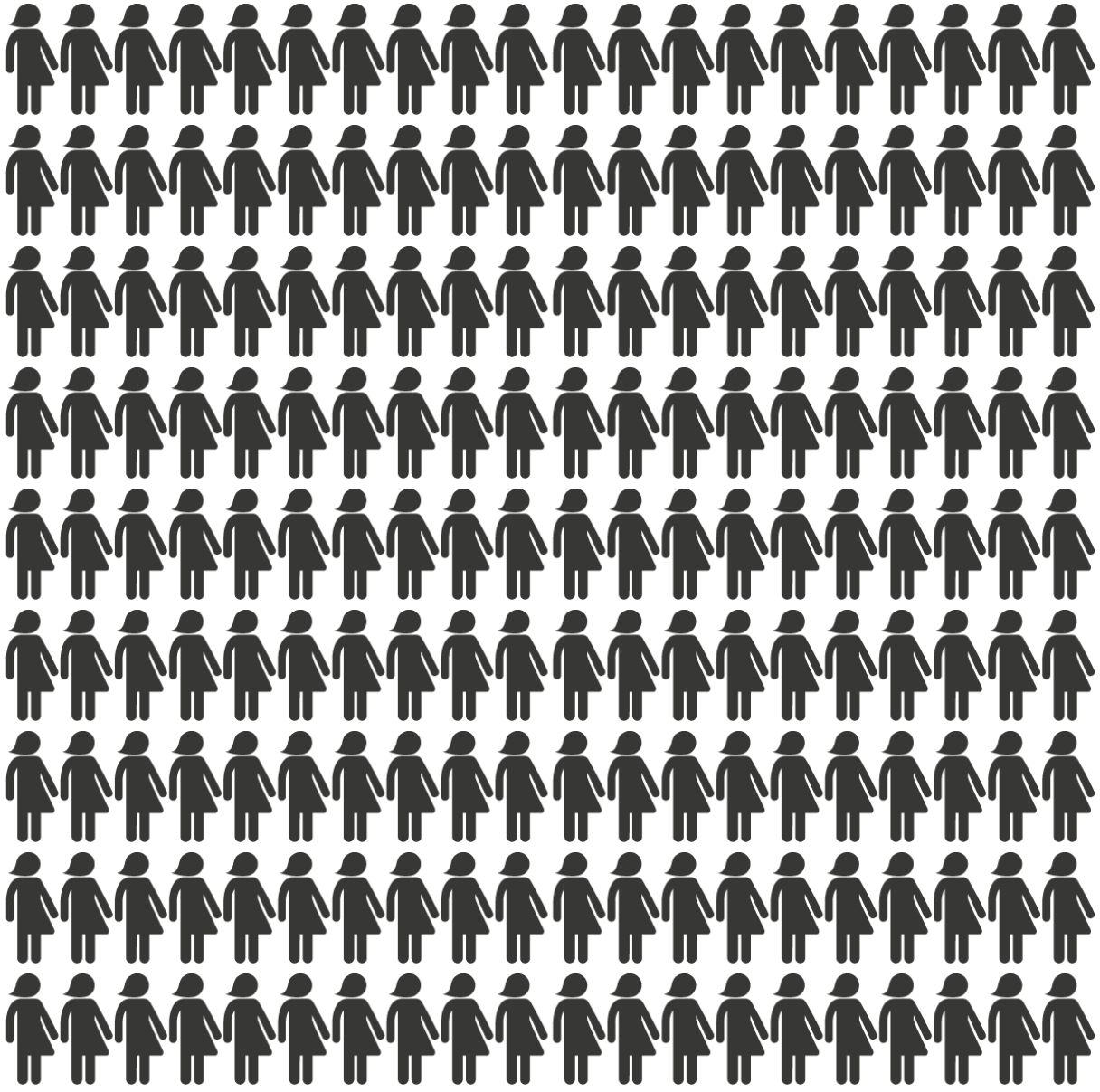
Of course, we can't know how long civilization will last. But what matters here is not any particular estimate: only that, on any reasonable guess, then, in expectation, the duration of future civilisation, and with it the number of future people, is immense.

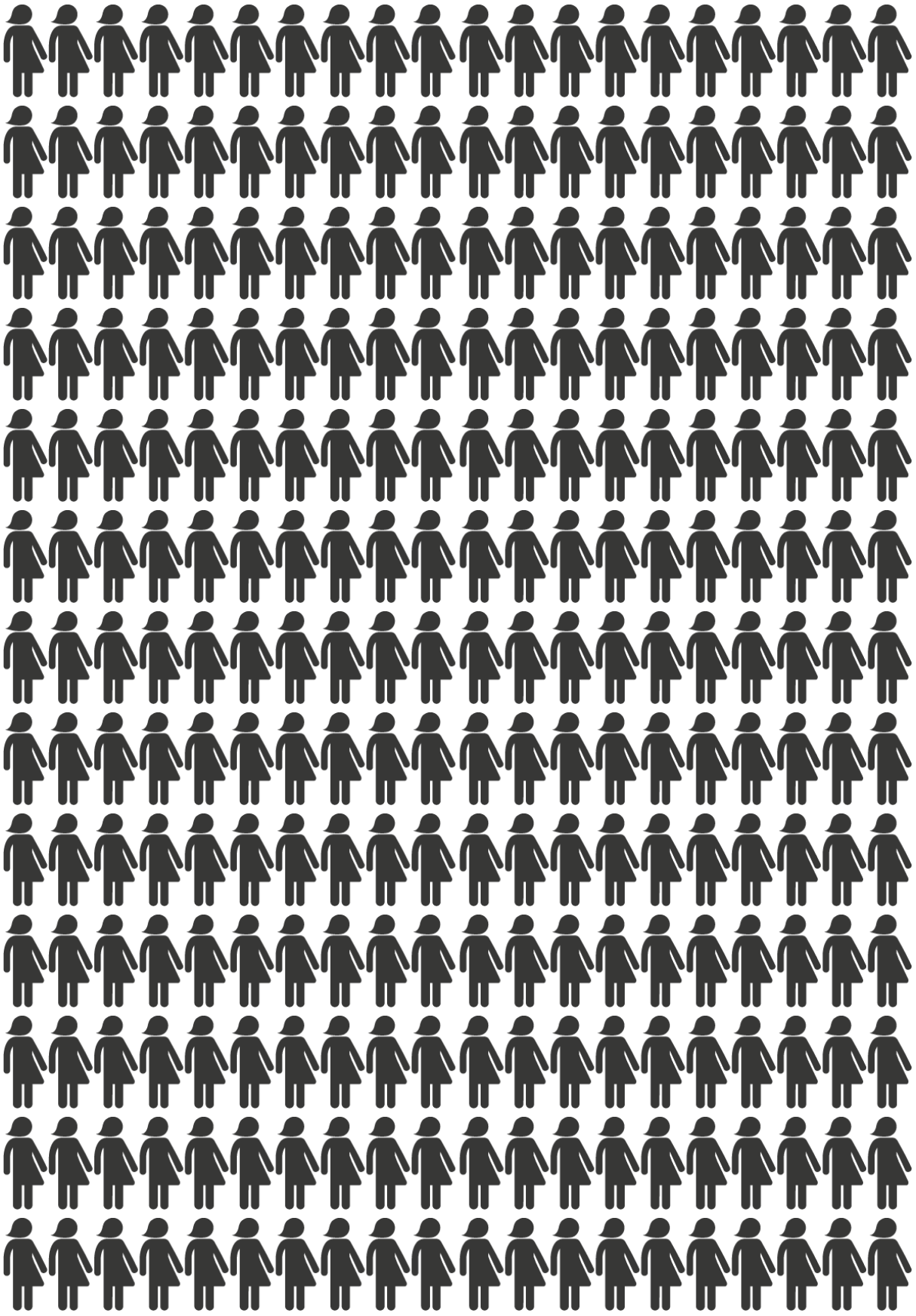
[28] [thank Toby for this latter fact]

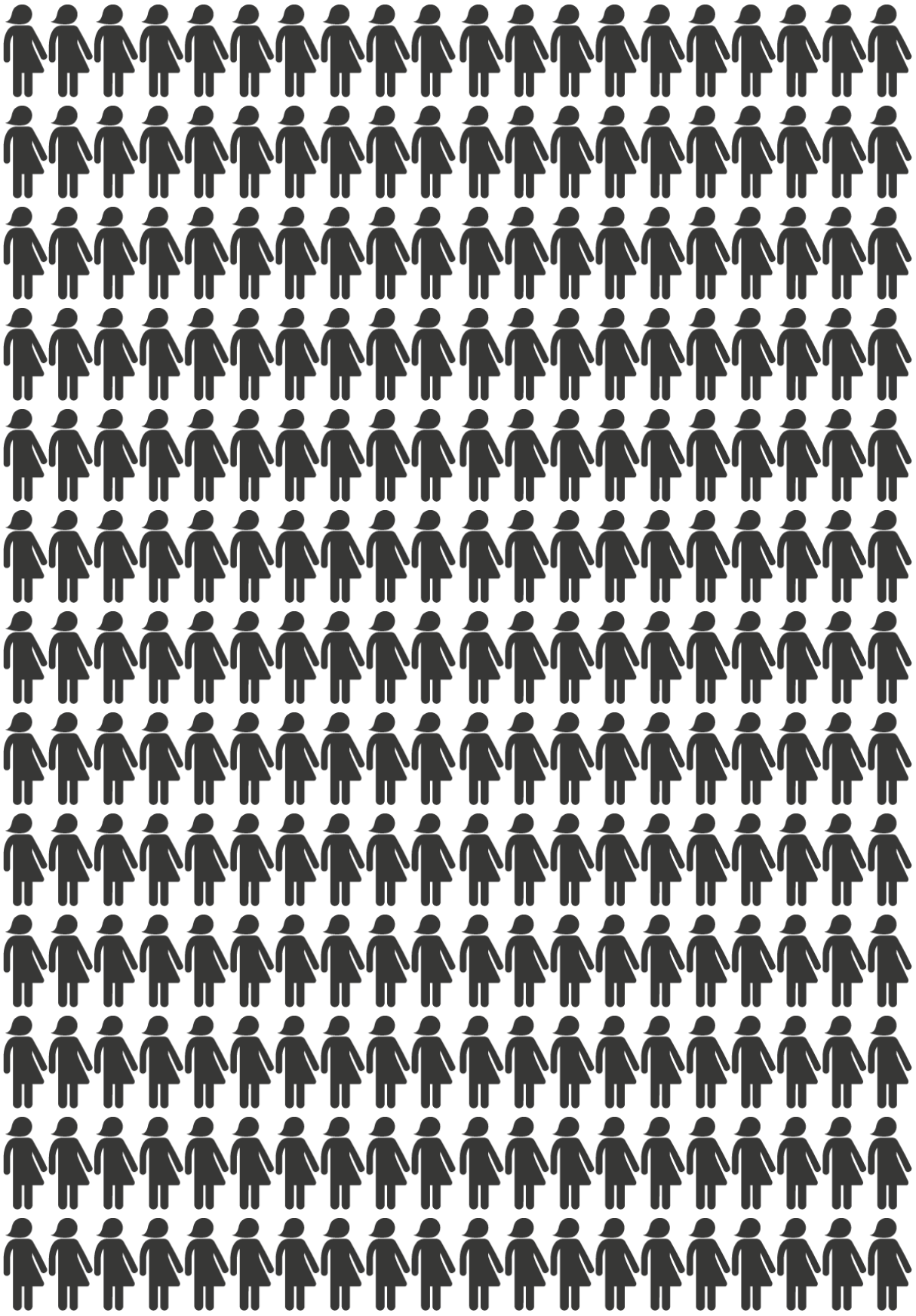
To appreciate the magnitude of this, take a look at the following diagram. Each stick figure represents ten billion people. Ten figures represent all those who have come before us, and the present generation is represented by a single person:

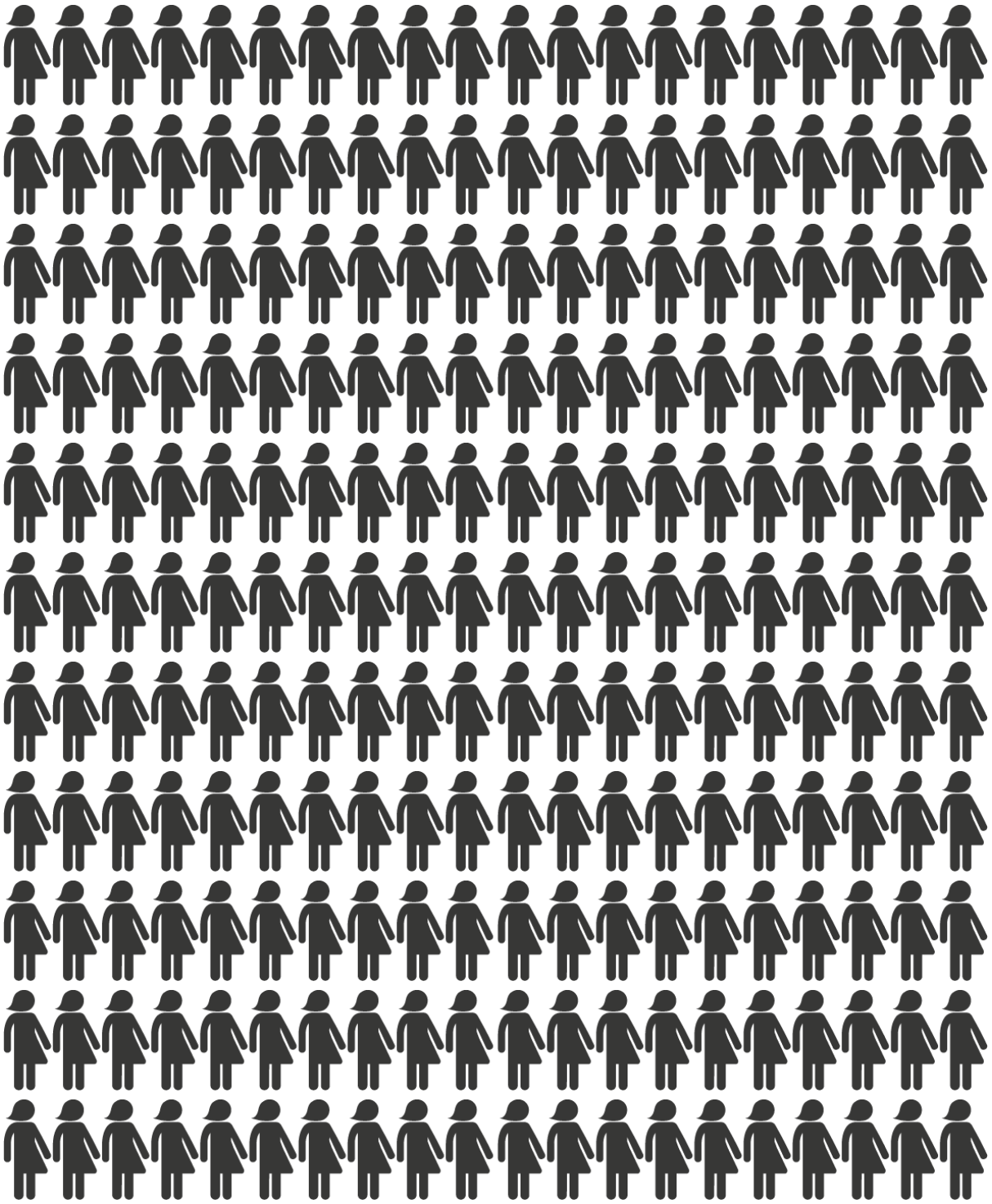


All the rest are those who are yet to come:









Represented visually, we begin to see the magnitude of what's at stake when we think about future generations. But the above diagram significantly underrepresents the scale we are talking about: I cut it short. To truly represent the number of future people in this diagram, I would have had to fill 4000 A4 pages with them — saturating this book ten times over.

This matters because, morally, the numbers matter. If you can save ten people or one person from dying in a fire, then, all else being equal, you should save the ten; if you can cure a

thousand people or a hundred people of a disease, you should cure the thousand. Every person is of equal moral value, so the importance of a given consequence scales proportionately with how many people are affected. So if future people outnumber us by thousands or millions to one, that means that their wellbeing is of immense importance.

Taking the sheer potential size of the future seriously requires a structural shift in how we think about our place in the world today. Currently, 'long-term' thinking might involve, at most, speculations about the next 30 years. When we are so rooted in what, on evolutionary timescales, is nanosecond-thinking, it can be hard to consider far longer time periods. But, if we truly care about the interests of future people, we have a duty to consider what the consequences of our actions might be over centuries, millenia, or even longer.

§4 The disenfranchisement of future generations

Future generations matter, and, because there are simply so many yet to come, they matter tremendously. But their wellbeing is not properly taken into account in the world today, because future generations are utterly disenfranchised.

That future generations are disenfranchised should be obvious given a moment's thought. They cannot represent themselves politically: they don't have a vote, and politicians therefore have scant incentive to consider how their policies impact the generations to come. Future people cannot lobby the government, and they can't bargain or trade with us, so they have no representation in the market, either. Even worse, future generations can't directly make their interests heard: they can't write articles about their lives to be published in the media; they can't march in the streets about the issues that will affect them the most.[29]

In this way, longtermism is an extension of the ideals of those previous social justice movements, such as civil rights and women's suffrage, which have sought to give greater influence to disempowered members of society. However, while past movements have removed barriers to disenfranchised groups, allowing the silenced to speak and be heard, we cannot do the same for future generations. It's up for us to represent their interests, because it's not possible for them to stand up for themselves.

---

[29] [felt too long, like a digression / bit random. But I find the point powerful.] [This fact becomes even starker when we compare future generations to one of the most disenfranchised groups in the current generation: the extreme poor — the 700 million people [check] globally who live on less than $2/day (where this is adjusted for purchasing power, i.e. it's the equivalent of what $2 could buy in the US today. [ref, make number precise.]) Political decisions made by rich countries regularly impact the lives of those in poor countries: [examples - whether to go to war with neighbours, the fairness (or not) of trade deals; international patent rights; tariffs; fossil fuel emissions; where to focus R&D efforts]. For this reason, the extreme poor regularly get a raw deal. But they can still make their voices heard to some extent: the conditions of extreme poverty can be represented, by themselves or others, through accounts in books and newspapers; the countries they live in have a vote at the UN. And this has meaningful consequences: though the concern for those in extreme poverty does not go nearly far enough, they are given at least some meaningful consideration within rich countries. Every rich democracy has a minister for international development [check]; there is an international standard that [rich countries] will spend 0.7% of their GDP [check] on international development, and [] countries are meeting that standard. There is nothing comparable for the political consideration of future people: a few countries, such as Wales [] have created a [minister or ombudsperson] for future generations, but their powers are extremely limited.]

You might wonder at this point whether there is any way, even in principle, that we could give political representation to future generations, and therefore whether discussion of their lack of political representation is idle. But that would be to misunderstand my point. I'm not claiming that there is any sense in which future generations 'should' get the vote. Rather, I'm pointing to the uncontroversial fact that they *don't*: and that therefore the normal democratic means by which different groups can exert political pressure and get treated more-or-less fairly aren't available to them. So, if there is some choice that the present generation can make, like whether to wantonly burn fossil fuels, which would benefit the present generation at the expense of future generations, we should expect the action that sacrifices the future to be taken every time. It is only the intergenerational altruism of some of those in the present that gives future generations any consideration at all; because of this, if it is possible to predictably impact the long-run future, we should not expect the best opportunities to have been taken.

You might also come away from this discussion feeling pessimistic. The civil rights and women's rights movements were successful, in significant part, because they were able to form a collective of the disempowered. Each individual Black person or woman at the time had little political power; but by coming together they had considerable clout. Such an option is not possible for future generations. So perhaps we have to accept that they will never get appropriate consideration, and we should give up hope.

But this pessimism would be too hasty. Though it is true that many social justice movements in the past have succeeded via members of the disenfranchised group taking collective action, not all have done so. For example, the environment cannot represent itself: it does not and cannot have a vote, and so is, like future generations, reliant on the moral concern of individuals in the present. Yet the environmentalist movement has had enormous growth and considerable success over the last 60 years.[30] The modern US environmentalist movement did not exist before the 1960s, and before then environmental issues were largely not on the radar of most people. But by the first Earth Day on April 22nd 1970, 20 million people — 10% of the US population at the time — participated in peaceful demonstrations to show support for environmental issues. The success of the environmentalist movement led to a plethora of new regulations over the 60s and 70s, including the National Environmental Policy Act and the formation of the Environmental Protection Agency, and led to international cooperation on environmental issues, too, such as the United Nations Environment Programme, the Montreal Protocol's ban of chlorofluorocarbons in 1987, and the formation of the Intergovernmental Panel on Climate Change. If concern for the environment can move from being fringe to thoroughly mainstream over the course of a century, purely through the efforts of altruistic individuals, so can concern for the welfare of future generations.

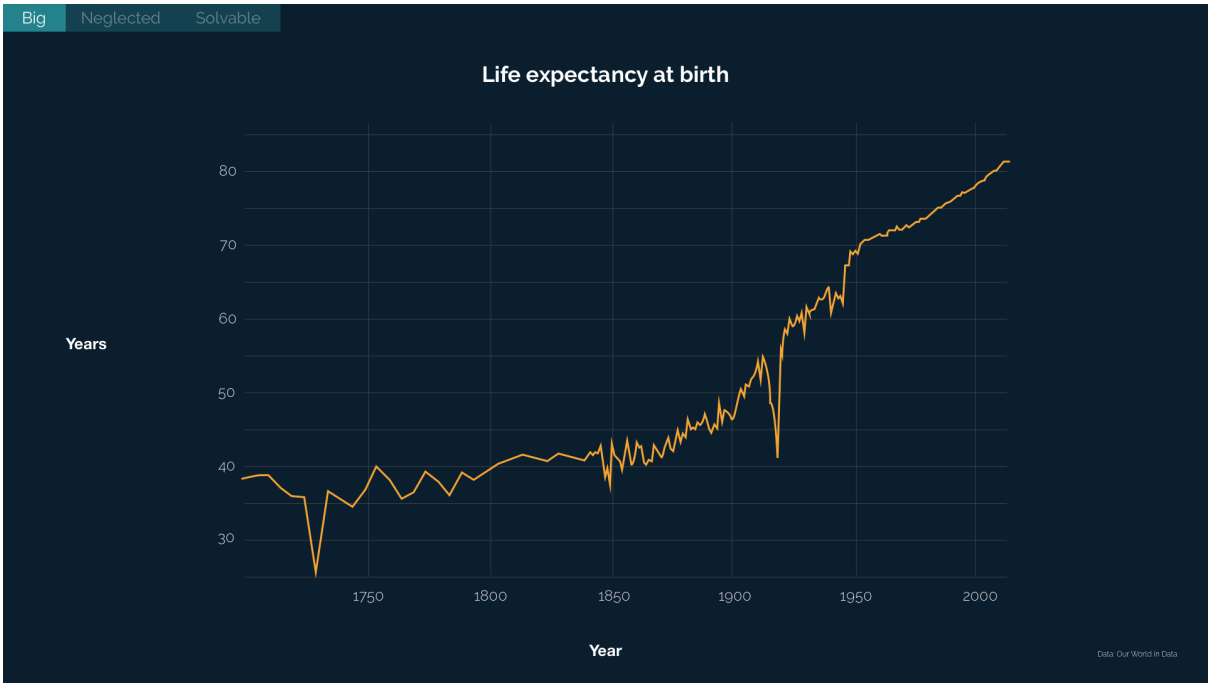
§5 A direct appeal to the importance of the future

The argument I've given so far shows, I believe, that longtermism is the natural consequence of taking seriously the same impartial, egalitarian ideals that motivated abolitionism, civil rights and women's rights. The weight of this argument is what first moved me to take

[30] Much from here: https://docs.google.com/document/d/19igUPkUqW1RwOdkLn7MVc-hiFF1CLuGxpnN0DbM9ZXA/edit, which itself cites Luke M.

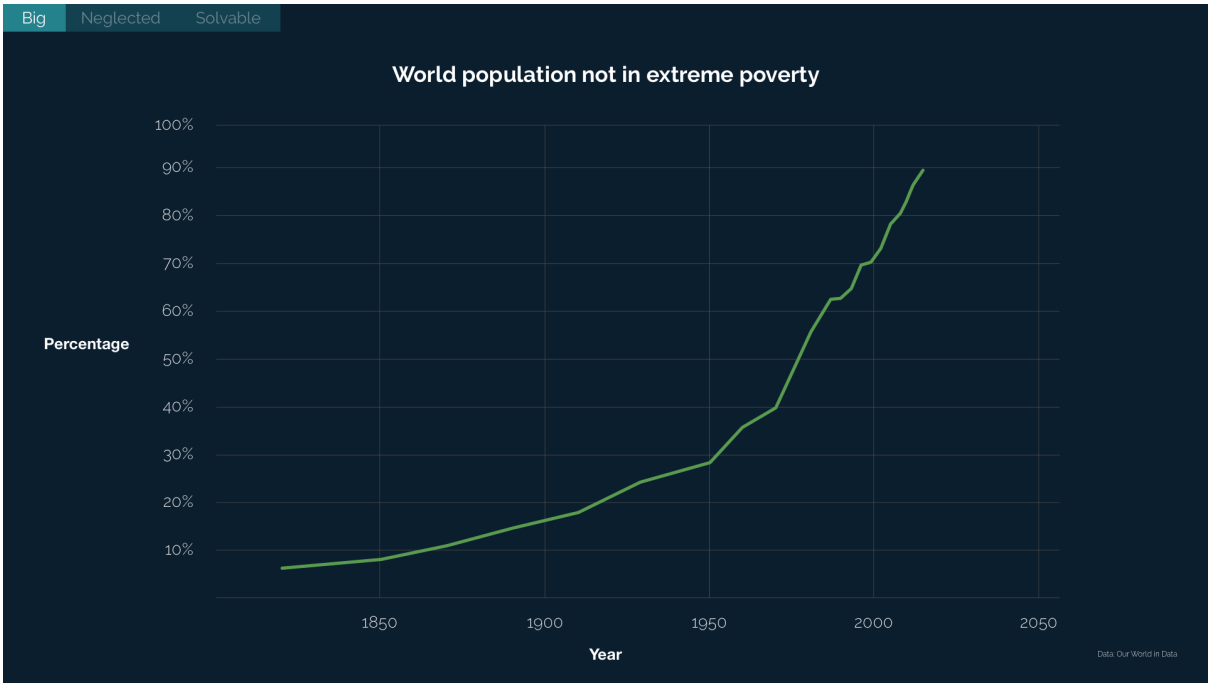
longtermism seriously. But now I find a straightforward and direct argument just as compelling, based on the simple idea that the future might be extraordinarily good or inordinately bad, and that we can affect the likelihood of which of these futures come about.

When thinking about how good the future could be, we should reflect on the progress that we have made over just the past two centuries. Given all the problems we currently face, it may not feel intuitive to think about the ways in which the world is getting better, but the past was a far more miserable place than we typically imagine, and in many ways things have improved substantially over time. For example, here is global life expectancy at birth:[31]

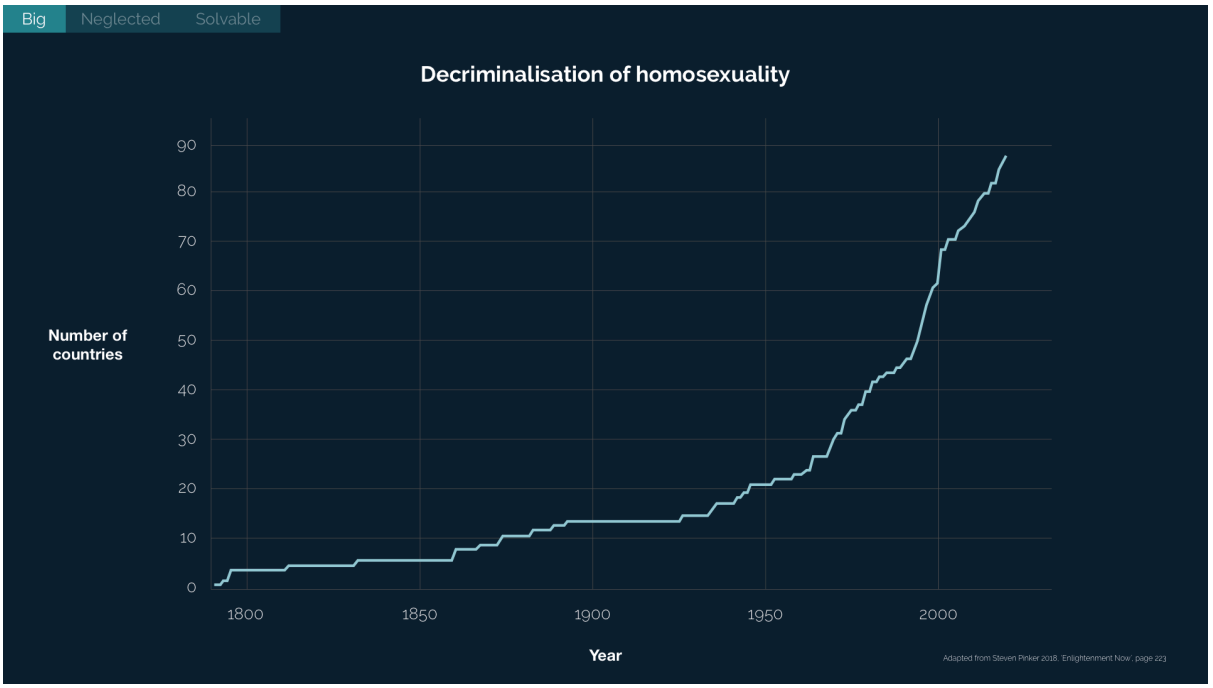


Here is the proportion of people worldwide who are *not* living in the most extreme poverty:

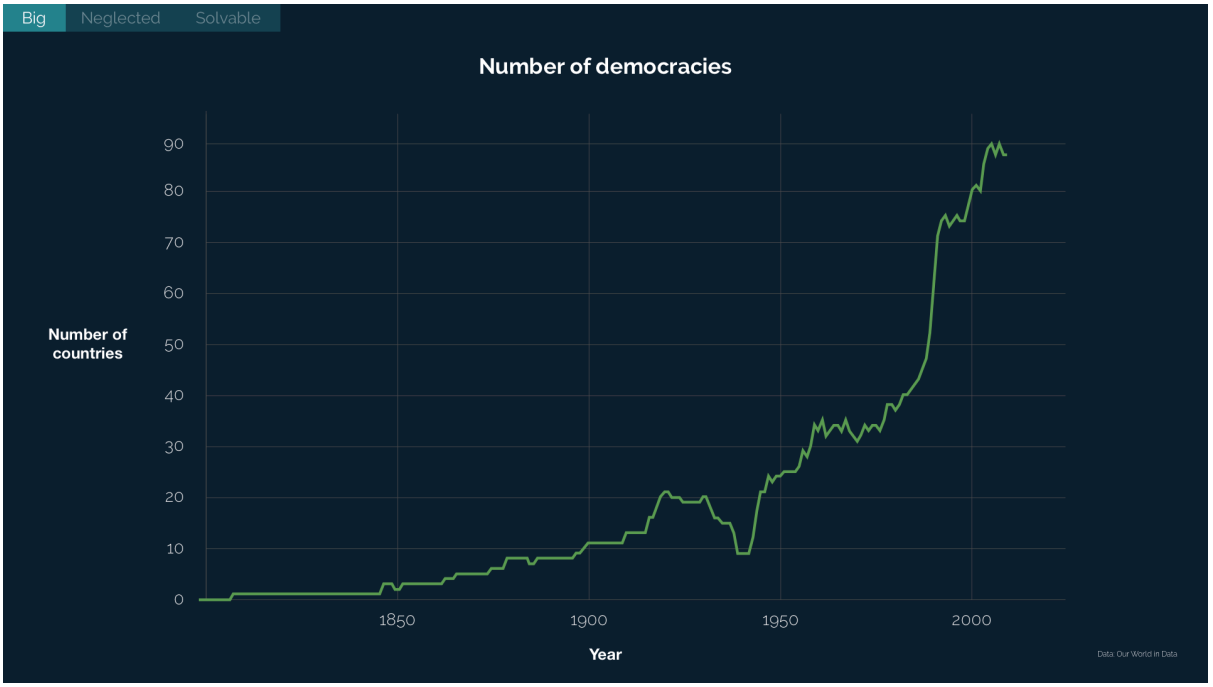
[31] All from Our World in Data



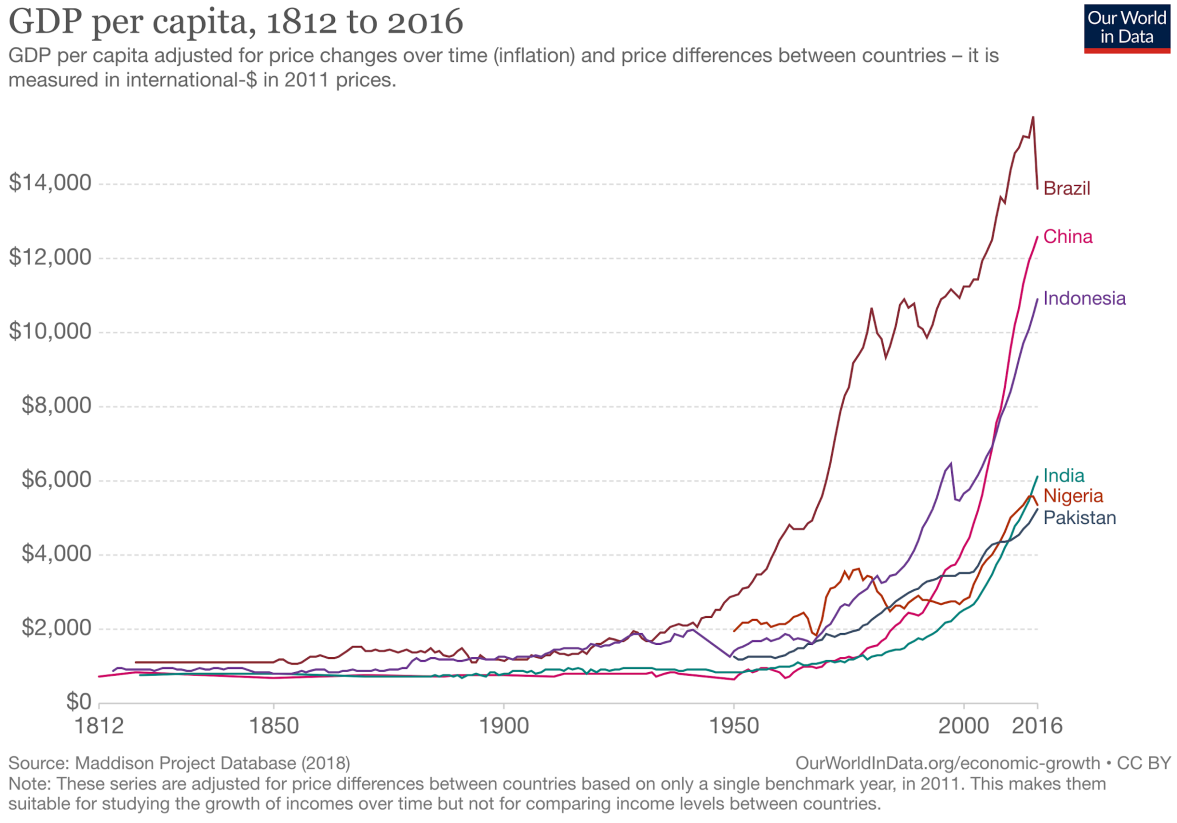
Here are the number of countries that have decriminalised homosexuality:



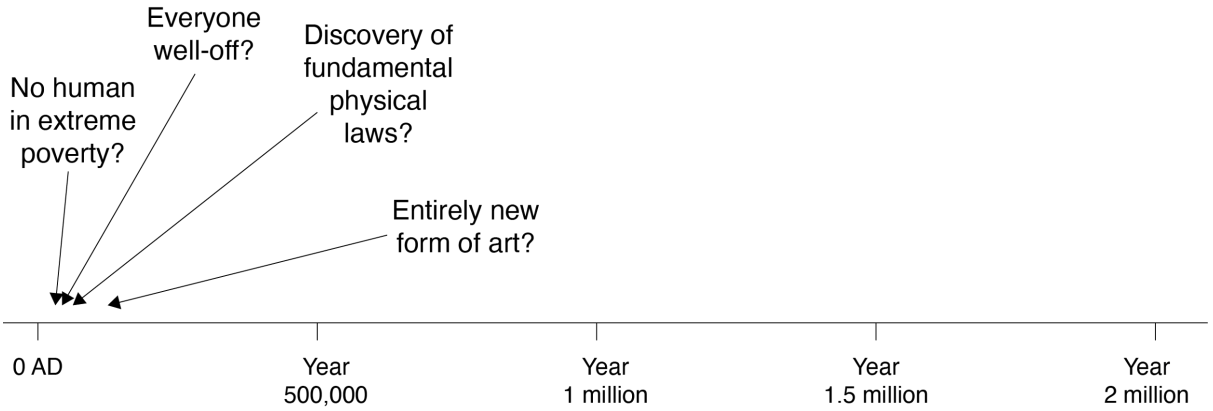
Here are the number of democracies over time:



Here is GDP per person for the six most populous poor and middle-income countries over time:



The world is still an unjust place, full of unnecessary suffering. We as a civilisation still have a long way to go. But the past was truly horrible, and in a wide variety of ways, the world has been getting better over the last two centuries.[32] If we achieve sustainable economic growth and technological progress, and the trends of the past 200 years continue, then within a century we should expect essentially no-one to live in extreme poverty; within 400 years we should expect everyone in the world to be better-off than the millionaires who are alive today. Beyond that point, we can only speculate. Perhaps in a few centuries we will have discovered the fundamental physical laws that govern our world, or invented entirely new forms of art — music that we lack the ears to hear. Given thousands of years of progress into the future, we could create eutopia.[33]



To get a sense of just how good this future could be, reflect on the very best moments of your life. For me, that's the intense love I've felt at the weddings of my best friends; the consonance of sight and sound at art and music festivals; the happy exhaustion I've felt at the top of a Scottish mountain. And now imagine if all of life were as good as those moments: no sadness, no suffering, no boredom, unless you choose to experience them for the variety; and imagine if everyone had access to such a life. That future is possible for us. Indeed, the future could be much better than that again: it would be surprising if the best human experiences were as good as it could get; future technology might allow us to experience heights of joy and wonder that are inaccessible to us today.

Such progress, however, is not inevitable. The Roman Empire at the turn of the first millennium, the Abbasid caliphate in the 8th to 10th centuries, and the Song dynasty in China in the 10th to 13th centuries are all examples of epochs of comparative economic flourishing and intellectual progress that preceded centuries of decay and stagnation.[34] Humanity's future could take the form of a gradual erosion of the progress we've made so far, or it could end

---

[32] Briefly discuss the ways in which I agree and disagree with Pinker (animals, war, environment; also caveat that we're making this judgment from the perspective of the dominant morality; many moral systems would have a very different view). Reference Max Roser's blogpost for what I see as the right framing. https://ourworldindata.org/much-better-awful-can-be-better

[33] Throughout, I use the term 'eutopia' which means "good place": an ideal state for society, as a practical aspiration, rather than 'utopia' which means "no-place". This is true to Thomas More's original coining; in his book, Utopia was not a representation of an ideal society. I also want to strongly differentiate the tentative, progressive and open-minded eutopian thinking that I advocate for later in the book from the often fanatical utopian movements of the past.

[34] Goldstone, Jack A. "Efflorescences and economic growth in world history: rethinking the" Rise of the West" and the Industrial Revolution." Journal of world history (2002): 323-389.

with a bang — a technologically-induced catastrophe, with the survivors eking out a miserable existence on a barely-habitable planet. Already we see some hints of the difficulties that continued technological development poses, via man-made climate change, the creation of potentially civilisation-ending weapons of mass destruction, and the increasing difficulty of furthering our scientific knowledge.

Moreover, even if we do succeed at generating long-run economic and technological development, this is worth nothing if the moral values that guide the world are mistaken. And it isn't by any means guaranteed that the values that shape the future will be good ones; even if we think that we've made moral progress to date, there are historical examples of moral retrogression. Europe, for example, saw a steady decline in the institution of slave-owning from the 11th century onwards, replacing it with serfdom, and by the 15th century there were few slaves on European soil. But this moral progress was quickly undone after Europe started colonising the New World, creating slavery's most despicable incarnation in the Atlantic slave trade. Similarly, the centuries that followed the Enlightenment saw a bloom in egalitarian moral views, but the fascist ideology that took over Germany in the early 20th century, premised on a hierarchy of races, sought to reverse that trend.

Future technological and political developments have the potential to make the worst-case outcomes for humanity, like slavery or totalitarianism, much worse again. Stalin, Mao and Hitler inflicted colossal misery on the world, each responsible for tens of millions of deaths, and for the suffering of hundreds of millions more. But if they had had access to radical life extension technology — which seems at least a possibility given the advances in biology we've seen in recent decades — they could have had the potential to hold onto power for thousands of years; if they had had access to cloning technology (which we now have), they could have groomed heirs, genetically identical to themselves, to continue their ideology indefinitely. With access to advanced genetic engineering, they could have created genetic hierarchies of workers; a Brave New World-inspired economy where people would be designed for the roles they will fill. With access to advanced surveillance technology and an automated police force and army, they could have guaranteed that any potential uprisings were quashed before they gained momentum. And if there had been a single world government, as both Hitler and Stalin desired, there would have been no external forces to restrict their power. Such scenarios might sound like science fiction, but surely they are no more outlandish than a description of contemporary society would seem to someone living just a few hundred years ago, for whom the ideas of the scientific method, the germ theory of disease, the theory of evolution, quantum mechanics and special relativity, modern democracy, limited liability companies, computers, cell phones, cars, the internet, flight, spaceflight, anaesthesia and countless more ideas and technologies that we nowadays take for granted would have been utterly alien.

Whether the future ends up as a eutopia or dystopia is up to us. It might seem that any attempt to try to affect which of these possible futures occur is quixotic in the extreme. That's a natural thought, but I think it's wrong. The fact that people are myopically concerned with what happens in the present day means that future generations' interests are neglected but it also means that we have remarkable leverage over the future. If you want to create a billion-dollar business, or make a killing on the stock market, you're competing against millions of other people who are trying to do the same. If you want to ensure the future goes as well as possible, you have very little competition: it's an open sea.

And the mere fact that we as individuals are small and the world is big does not mean that we cannot have leverage over the direction the world goes. The activism of a small group of moral radicals within the Quakers in the 1700s, over the course of a century, led to the abolition of slavery, the most repugnant institution man has ever devised. The US constitution was the result of debates among just 55 delegates at the 1787 Philadelphia Convention, arguing amongst themselves over the course of just four months, and it constrains the most powerful country on Earth to this day. The early Christians — Jesus, his disciples, and Paul the Apostle — laid down a worldview and a moral code that has guided the behaviour of tens of billions of people throughout history; similar can be said for Moses, Mohammed, Buddha, and Confucius. There are countless instances where individuals profoundly shaped the lives of future generations for centuries after their deaths. We have the opportunity to do the same.

We can think of civilization as a massive ocean liner bound on a long voyage. It might seem impossible for any one person to affect its path. But the direction of a ship can be altered by the comparatively minor force of its rudder, and over time a small difference in its direction can make an enormous difference to its trajectory. If that ocean liner left from London, the force of just one swimmer pushing on the rear of it, rotating it over the course of one day, would make the difference between its ending up in New York or Venezuela.[35]

§6 Summing up

The potential future of civilisation is vast, and the number of people who are yet to come is colossal. Because future generations count, morally, in just the same way that those of us in the present generation do, the interests of our descendants matter enormously. But, despite the importance of their wellbeing, future generations are utterly disenfranchised in the world today. This means that anything we can do to positively steer the long-run future is of tremendous moral importance. The future could be extraordinarily good or inordinately bad, and we have the power to influence which.

However, if you're like most people, I suspect that there is a line of objection to my argument that is top of your mind: that we cannot reliably predict the future, and that it's absurd to think we could take actions that shape the centuries or millennia to come, much less the entire course of future human history. The next chapter will take this objection head on, and discuss at length how we can predictably increase the chance of the future going well.

[35] [I got this estimate from a Facebook discussion (from physicists); need to make assumptions explicit.]