UNIVERSITY OF CAMBRIDGE
Faculty of Economics

Institute for New Economic Thinking

Cambridge-INET Institute

Cambridge-INET Working Paper Series No: 2016/19

Cambridge Working Paper Economics: 1660

BIRTH AND DEATH

Partha Dasgupta
(University of Cambridge)

It has long been known that in finite economies Classical Utilitarianism commends policies that encourage large populations. It has been known also that the stronger is the aversion to risk and inequality in the standard of living, the lower is the optimum living standard, and that the latter tends in the limit to Sidgwick's "hedonistic zero". A version of that extreme feature of the theory was subsequently named the Repugnant Conclusion (RC). Most escape routes from RC have invoked the language of "gains" and "losses", which are familiar notions in social cost-benefit analysis. Those notions have been found to lead to paradoxes involving the Non-Identity Problem. In this paper I start with Sidgwick's theory in its pristine form - the criterion for evaluating states of affairs is the sum of personal utilities - but recast it in a contemporary language: the ground of binding reason is taken to be "well-being", not "happiness", nor "agreeable consciousness". Sidgwick erred in his interpretation of the hedonistic zero, which may explain why the seeming pro-natalism inherent in his theory has been found to be repugnant by philosophers. Problems with Sidgwick's Utilitarianism lie elsewhere. An example is presented which invites an additional but relatively mild notion of person-hood into any theory that says that personal well-beings should be the sole basis for ranking states of affairs. A weak version of Agent-Relative theories is drawn from the example, which in the context of population ethics may be called Generation-Relative Utilitarianism. It has however been suggested that the theory is incoherent because it does not yield a binary relation between states of affairs. I show that the incoherence would arise only if states of affair were to be evaluated from nowhere, and that it is an essential feature of Generation-Relative Utilitarianism that the state of affairs from which other states of affairs are viewed, matters. The theory is then put to work in a model economy facing an indefinite future and a finite flow of resources. Empirical studies of Earth's life support system are then used to justify the choice of the model. I make use of contemporary global statistics to get a feel for the theory's implications for both population size and the standard of living. Population size is found to be smaller and the living standard higher than they would be under Sidgwick's Utilitarianism. Population ethics is then used to understand the nature of loss that would be suffered in the face of human extinction. Generation-Relative Utilitarianism is shown to arrive at the view that each generation is a trustee of the capital it inherits from its predecessor.

Birth and Death

by

Partha Dasgupta*

University of Cambridge
and
New College of the Humanities, London

First Version: April 2016
First Revision: June 2016
Second Revision: October 2016

* The author is the Frank Ramsey Professor Emeritus of Economics at the University of Cambridge, Fellow of St John's College, Cambridge, and Visiting Professor at the New College of the Humanities, London.

E-mail: partha.dasgupta@econ.cam.ac.uk

Acknowledgements

This article is an adaptation of Chapter 10 of *Time and the Generations*, a book I am preparing around my Arrow Lectures at Stanford University (1997), Columbia University (2011), and the Hebrew University of Jerusalem (2012); the 2000 Tanner Lectures on Human Values at the University of Michigan, Ann Arbor; and the 2011 Munich Lectures in Economics. Valuing potential lives and the related idea of optimum population have intrigued me ever since I was a graduate student, and I am grateful to the late James Meade for arousing my interest in them. The subject is hard; so hard that over the years I have fumbled about to find ways to express my disquiet with the dominant formulation of the problem and the literature surrounding the paradoxes it harbours. And I am all too conscious that readers may find me fumbling even now. Parental desires and needs in the face of socio-economic and ecological constraints are the basis on which economic demography has been built. Moral philosophers in contrast study population ethics, but shy away from characterising the constraints under which the ethics is to be put to work. No system of ethics should be expected to yield unquestionable directives in all conceivable circumstances, even to the same person. If we are to arrive at satisfactory policies, a suitable accommodation has to be found for the demographer's predictions, the economist's concerns, the environmental scientist's predilections, and the philosopher's sensibilities. These are early days even to attempt an integrated study. So I go about matters in a piecemeal fashion here. My idea is to construct a structure for population ethics that accommodates contemporary sensibilities over birth and death and then to put it to work on contemporary data on humanity's reliance on the biosphere.

The present essay builds on Dasgupta (1969, 1974a, 1989, 1994, 2005), a decadal engagement with population ethics. Over the years I have learned much on the subject from discussions with Geir Acheim, Scott Barrett, Simon Blackburn, Charles Blackorby, John Bongaarts, Christopher Cowie, Aisha Dasgupta, Shamik Dasgupta, David Donaldson, Paul Ehrlich, Nicholas Humphrey, Simon Levin, Karl-Goran Maler, Eric Maskin, James Mirrlees, Edmund Phelps, Peter Raven, Martin Rees, Paul Seabright, Amartya Sen, Robert Solow, and Menahem Yaari. I have also been much stimulated by the criticisms of my writings on population ethics by John Broome (Broome, 1996, 1999, 2004).

Sections 1 and 9 have been drawn from Dasgupta and Dasgupta (2016) and I am most grateful to Aisha Dasgupta for the many discussions we have had while preparing our paper on the place of reproductive rights in population ethics. An earlier version of this paper was prepared for a lecture on 14 January 2016 at a conference on Discounting and Climate Change held at the University of Oxford. For their comments I am grateful to Paul Kelleher, Kian Mintz-Woo, Matthew Rendall, and Dominic Roser. I have benefited greatly from the comments on previous versions by Simon Blackburn, John Broome, Christopher Cowie, Naomi Goulder, Rae Langton, Cheryl Misak, Stefano Zamagni, and in particular the insights of Simon Beard, Shamik Dasgupta, Ingmar Schumacher, Itai Sher, and Robert Solow.

My greatest debt is to Kenneth Arrow, discussions and collaboration with whom have regularly clarified my understanding of welfare economics and population ethics. This essay is dedicated to him.

Abstract

It has long been known that in finite economies Classical Utilitarianism commends policies that encourage large populations. It has been known also that the stronger is the aversion to risk and inequality in the standard of living, the lower is the optimum living standard, and that the latter tends in the limit to Sidgwick's "hedonistic zero". A version of that extreme feature of the theory was subsequently named the Repugnant Conclusion (RC). Most escape routes from RC have invoked the language of "gains" and "losses", which are familiar notions in social cost-benefit analysis. Those notions have been found to lead to paradoxes involving the Non-Identity Problem. In this paper I start with Sidgwick's theory in its pristine form - the criterion for evaluating states of affairs is the sum of personal utilities - but recast it in a contemporary language: the ground of binding reason is taken to be "well-being", not "happiness", nor "agreeable consciousness". Sidgwick erred in his interpretation of the hedonistic zero, which may explain why the seeming pro-natalism inherent in his theory has been found to be repugnant by philosophers. Problems with Sidgwick's Utilitarianism lie elsewhere. An example is presented which invites an additional but relatively mild notion of person-hood into any theory that says that personal well-beings should be the sole basis for ranking states of affairs. A weak version of Agent-Relative theories is drawn from the example, which in the context of population ethics may be called Generation-Relative Utilitarianism. It has however been suggested that the theory is incoherent because it does not yield a binary relation between states of affairs. I show that the incoherence would arise only if states of affair were to be evaluated from nowhere, and that it is an essential feature of Generation-Relative Utilitarianism that the state of affairs from which other states of affairs are viewed, matters. The theory is then put to work in a model economy facing an indefinite future and a finite flow of resources. Empirical studies of Earth's life support system are then used to justify the choice of the model. I make use of contemporary global statistics to get a feel for the theory's implications for both population size and the standard of living. Population size is found to be smaller and the living standard higher than they would be under Sidgwick's Utilitarianism. Population ethics is then used to understand the nature of loss that would be suffered in the face of human extinction. Generation-Relative Utilitarianism is shown to arrive at the view that each generation is a trustee of the capital it inherits from its predecessor.

Contents

People have children for many reasons. The mix of motivations depends on the customs and institutions we inherit, as well as on our character and circumstances. That children are valuable in themselves is emotionally so compelling that it may seem too obvious to require acknowledgement, but social anthropologists have shown that children are not just valuable to us because of the innate desire we have to bear and rear them, but also because they represent the fulfilment of tradition and religious dictates, and because they are the clearest avenue open to self-transcendence. A common refrain, that our children are priceless, is an expression of how innately valuable they are to us.[1]

Nevertheless, in places where formal institutions are underdeveloped, children also substitute for other assets, and are thus also valuable for the many benefits they bring to their parents. This is most apparent in the poorest regions of the world. Children serve as security in old-age in places that have neither pension schemes nor adequate land markets. They are also a source of labour in households possessing few labour saving devices. Children mind their siblings, tend to domestic animals, pick berries and herbs, collect firewood, draw water, and help with cooking. Children in poor countries are valued by their parents also as capital and producer goods.[2]

1 Economic Demography

Those childhood activities are so unfamiliar today in the West that they direct us to study the mix of motivations governing procreation by contrasting rich regions from poor regions. There are notable exceptions of course, but broadly speaking fertility and mortality rates are high and health status and education attainments are low in poor countries, whereas the corresponding statistics in rich countries read the other way. Table 1, which presents a snap shot (year 2010), speaks to that by displaying the classification used at the World Bank, where "rich" and "poor" are defined in terms of GDP per capita. Countries have been known to make a transition from one category to the other (that's what economic development is usually taken to mean); moreover, the bulk of the world's population and a majority of the world's poorest people live in neither rich nor poor countries, and international statistics say there are enormously rich people in poor countries. It nevertheless pays to study sharp contrasts, as in Table 1.[3]

---

[1] One such injunction emanates from the cult of the ancestor, which takes religion to be the act of reproducing the lineage. See Fortes (1978).

[2] In South Asia children have been observed to be at work from as early an age as six.

[3] Evolutionary biologists distinguish reproductive strategies in the animal kingdom that correspond to the two extreme types of fertility and human capital outcomes displayed in Table 1. An *r*-strategy gives rise to many offspring combined with low parental investment, and each offspring has a low probability of survival. In contrast a *K*-strategy gives rise to few offspring over a longer lifespan and high parental investment. The motivation driving animals is inclusive fitness. Each of the strategies is an equilibrium, depending on the underlying evolutionary game. I am grateful to Nicholas Humphrey for discussions on the significance of animal behaviour for

Reproductive decisions have consequences for others (including our descendants) that are unaccounted for under the terms of prevailing institutional arrangements (markets, government policy, communitarian engagements, religious injunctions). Economists use the term "externalities" to denote those consequences of our decisions for others that are unaccounted for. Externalities raise deep ethical issues. Not only do they extend to contemporaries and can be expected to extend to future people, it is also that some people will be born in consequence of the decisions we take, while some who would have been born had we acted otherwise will not be born.[4]

Caldwell (1981, 1982) drew on a suggestion implicit in the contrast displayed in Table 1; that the intergenerational transfer of economic resources is from children to parents in poor societies, but from parents to children in rich societies. The suggestion has been easier to confirm in rich countries, where the rate of investment in children's education has been found to be as high as 6-7 per cent of GDP (Haveman and Wolfe, 1995). Confirming the reverse flow in poor countries has been a lot harder, in part because data are sparse but in part also because even within poor regions there are significant differences in attitudes toward reproduction. Those differences are traceable to kinship structures, marriage practices, and rules of inheritance. The implied line of thinking says that over the long run it is differences in institutions and social norms - originating perhaps in some measure in geography - that are the reasons behind differences in reproductive behaviour among groups. Theoretical models have been built on that premise. Causality isn't traced to differences in income or wealth. It is not that fertility and mortality rates are high and health status and education attainments are low in poor regions *because* people there are poor, it is that very low incomes go hand in hand with those features of life. The variables are mutually determined over time.[5]

Table 1 is a snapshot. There is no suggestion that poor regions will remain poor, nor that

---

population ethics.

[4] Decisions on consumption, production, and our use of the natural environment also create externalities. They are connected to reproductive externalities. Here I focus on decisions over reproduction. We uncover the pervasiveness of externalities in Section 9.

[5] For theoretical models that speak to the mutual determination, see Dasgupta (1993, 2010), Brander and Taylor (1998), Harford (1998), and Dasgupta and Ehrlich (2013). Sub-Saharan Africa has long been regarded as special, even among poor regions (Goody, 1976; Bledsoe, 1994; Guyer, 1994; Bongaarts and Casterline, 2013). In an early review of fertility intentions Cochrane and Farid (1989) noted that both the urban and rural, the educated and uneducated in sub-Saharan Africa have more, and want more, children than their counterparts in other less-developed regions. Even young women there expressed a desire for an average of 2.6 more children than women in the Middle East, 2.8 more than women in North Africa, and 3.6 to 3.7 more than women in Latin America and Asia. Updated versions of these figures are available, but it is worth considering the data from the mid-1980s because the income gap between Africa and the rest of the developing world was smaller at that time than it is now.

rich regions may not find their place reversed in the long run. Regional differences in fertility, education, and output per capita were slight until the start of the Early Modern era (roughly, 1500 CE). Global aggregates of earlier eras look much the same as their regional aggregates.[6] Although regional aggregates have diverged since then, global aggregates (a weighted average of regional statistics) have shown a steady move toward and beyond the "fertility transition"; that is, the transition from high fertility and mortality rates to low fertility and mortality rates.

Economists have offered a number of explanations for the historical experience.[7] What is common to them is a presumption that parental choices over fertility, consumption, and investment determine long run outcomes. The models trace the relative urgencies of parental needs, desires, and obligations to the constraints on choices facing parents in each generation. Some authors stress economic constraints, others pay attention to social and ecological constraints. Economic demographers have commonly avoided moral theories in their study of reproductive behaviour.[8]

In contrast, philosophical discourses on population have been built exclusively on normative reasoning, directed at (broadly speaking) four questions: (1) What are the nature, ground, and limits of parental responsibility for existing children? (2) Does producing a child violate the rights or interests of children the couple already have? (3) Do individuals have a duty not to have children whose lives may be considered "bad"? Do they have a duty to have children whose lives might be considered "good"? (4) What procedure should be used to determine the relative value of possible populations so as to decide which of those would be "best"? In addressing these questions, however, moral philosophers have avoided offering even a sketch of the constraints under which that reasoning is to be put to work.

One way to contrast these two disciplinary approaches is to say that the economic demographer's task is to explain Table 1 and that the aim in population ethics is to produce a normative theory that one could use to evaluate the behaviour patterns that give rise to Table 1 and then prescribe better ones. That said, no system of ethics should have to yield incontrovertible directives in all conceivable circumstances. In order to arrive at satisfactory demographic policies, a suitable accommodation has to be found for the demographer's evidence, the economist's concerns, the environmental scientist's predilections, and the philosopher's sensibilities. In the remainder of the paper I study population ethics with an eye on ecological constraints. We will find that it is possible to produce sensible answers to problems that arise in

---

[6] The classic on this is Maddison (2001), who provided estimates of expectancy of life at birth, population size, and output from 1 CE until year 1998 in various regions of the world.

[7] See the references in footnote 5, and Galor (2011).

[8] "Parental choices" are only a contemporary way of alluding to choices that are often, perhaps even usually, influenced by the extended family, kinship, and the couple's peer group. On this see Dasgupta (1993, 2010) and Dasgupta and Dasgupta (2010).

the world we have come to know even if we are unable to make progress in answering the kinds of question population ethicists usually ask in the context of extreme cases.

2 Utilitarian Ethics

In his introduction to Utilitarianism, Sidgwick (1907: 415-416) wrote:

"... if we take Utilitarianism to prescribe, as the ultimate end of action, happiness as a whole, and not any individual's happiness, unless considered as an element of the whole, it would follow that, if the additional population enjoy on the whole positive happiness, we ought to weigh the amount of happiness gained by the extra number against the amount lost by the remainder. So that, strictly conceived, the point up to which, on Utilitarian principles, population ought to be encouraged to increase, is not that at which average happiness is the greatest possible ... but that at which the product formed by multiplying the number of persons living into the amount of average happiness reaches its maximum."

Scholars have found a problem with the opening sentence. Sidgwick starts by noting that the ground of binding reason is total happiness ("happiness as a whole") but goes on to appeal to gains and losses when applying the Utilitarian calculus. That the two aren't equivalent if the timing of conceptions (never mind the number of conceptions) is affected by decisions was noted by Parfit (1976), Schwartz (1978), and Adams (1979) and shown to give rise to what is now known as the Non-Identity Problem.[9] The problem, as Parfit (1982, 1984) posed it, goes like this:

"A woman suffers from a medical condition. There is a large chance that if she were to conceive now, the child would suffer from a disability but would otherwise enjoy a good life. However, a minor medical treatment would cure the woman within a month. Once cured, any child she bears will be free of the disability and enjoy a life of high quality. The woman is somewhat impatient to conceive. Ought she to wait a month, or would it be reasonable of her to conceive now?"

One can argue that it is reasonable she conceives now. After all, or so the argument could go, the woman's feelings matter and the child she conceives now can't complain later that she was unfair to him, that she should have waited and undergone the medical treatment, that had she done so he would have had a better life. The reason he would not have grounds for complaint is that had she waited, the child she would have conceived wouldn't have been him. Nor, or so the argument may continue, can some unconceived child complain that the woman prevented him from being born by being hasty. Nonetheless, as Parfit and others observed, there is a strong intuition that the woman should wait. And the intuition is built not only on the thought that good lives are an intrinsic good, but also that better lives are intrinsically even better.

In this intuition the basis for evaluation is not "gains" and "losses", but total utility. And an ethics grounded on the latter reflects the view that lives have an intrinsic value of their own,

[9] I have taken the attribution from the essay by Kavka (1982) which, with an accompanying essay by Parfit (1982), explored the complexities of the Non-Identity Problem.

supporting Sidgwick's insistence that it is "happiness as a whole" that is the ultimate basis for population ethics. Utility is taken to be a numerical measure of happiness. The Non-Identity Problem doesn't rear its head in the version of Sidgwick's Utilitarianism I work with because it asks us to evaluate alternative states of affairs in terms of the sum of personal utilities. State of affairs *X* is judged to be superior to state of affairs *Y* if total utility in *X* exceeds total utility in *Y*. That comparison can be made without invoking benefits and losses in utilities.[10]

Sidgwick's theory, or what is widely known as Classical Utilitarianism, involves two related notions: a person's happiness (whose numerical measure is her utility) and the way individual utilities are combined to yield a notion of aggregate utility. The aggregate he commended is obtained by summing individual utilities. Subsequently I will depart from Sidgwick on the latter count and propose an alternative that commends itself. Sidgwick (1907: 119-150) contains three chapters on empirical hedonism, where the sense in which "happiness" (or "agreeable consciousness") is used is a lot more considered than is suggested in the frequent criticism that Classical Utilitarianism views humans to be mere pleasure machines. Nevertheless I am reluctant to give the impression that the ethical structure I am led to relies on Sidgwick's notion of the personal good, namely, happiness. So I shall use the term "well-being".[11]

3 Plan of the Paper

"Well-beingism" reads distinctly odd. So we will call the theory that evaluates states of affair in terms of the sum of individual well-beings, Total Utilitarianism, even though the raw material out of which it is constructed is personal well-being, not personal happiness. Section 4 explores how far that theory is able to guide population ethics before running into trouble. I do that by putting the theory through its paces in a timeless world with finite resources.

My idea here is to construct a framework for population ethics that accommodates contemporary sensibilities over birth and death. In Section 5 I show that Sidgwick erred in his reading of a "neutral life", which is a life that goes neither well nor badly. That error recurs in contemporary writings on population ethics. Criticisms of Total Utilitarianism built on that misreading and the paradoxes the theory is thought to harbour are thus misplaced. In Section 6 an example is constructed to show that problems with Total Utilitarianism lie elsewhere. I explore a way that suggests itself for amending the theory. It involves introducing a notion of person-hood and combining personal well-beings in a different manner from Total Utilitarianism. The amended version evaluates states of affair in terms of a suitably *weighted* sum of individual

[10] There has been a revival of interest in the Non-Identity Problem. Boonin (2014) is a book-length treatment and has a comprehensive set of references. I have gained much from reading an unpublished paper on the subject by Christopher Cowie (Cowie, 2015).

[11] Griffin (1986) contains a measured, book-length analysis of the concept in its many guises. To put it in the briefest way, he thinks of personal well-being as the extent to which one's informed desires are realized. Griffin also discusses the problems that arise in measuring the object.

well-beings, in which the weight awarded to well-beings of potential people is lower than to the well-beings of present and future people. I call the amended version Generation-Relative Utilitarianism. In Section 7 Generation-Relative Utilitarianism is applied to a simple example in a timeless world. The theory is shown to recommend a smaller population and a higher living standard than Total Utilitarianism. Parfit (1976) and, following him Broome (2004: 157-162), have argued that Generation-Relative ethics is incoherent because it doesn't yield a binary relation between states of affairs. I show that the theory only appears incoherent, and does so because the the theory Parfit appeals to views states of affairs from nowhere. Generation-Relative Utilitarianism instead evaluates states of affair from inside the states of affairs themselves.

In Sections 8-10 the theory is applied to a world that moves through time. So as to allow for easy comparison with the timeless world of Sections 4 and 7, the model I construct does not permit capital accumulation. We assume that nature provides a constant flow of consumption services, like manna from heaven. Assuming that people are identical, consumption at each date is distributed equally among all who are present. The ethical tradeoff is then between population size and the average level of consumption. We find that the ideal, or optimum, policy under Generation-Relative Utilitarianism can only be achieved if there is an implicit understanding among the generations. That understanding sees each generation as a trustee of the basis on which the optimum policy is implemented.

An absence of production in our model will be shown not to mislead.[12] What could mislead however is the presumption that the economy's institutions are so good that we are wholly discouraged from wasting nature's services (in Section 9 I review global statistics that are pertinent to the issue). But it is the natural assumption to make about a world that seeks to identify an ideal population policy.

In Section 10 I work with the model by using the Global Footprint Network's estimates of the magnitude of the ecological services our biosphere offers us. I conduct a sensitivity analysis of Generation-Relative Utilitarianism by varying key ethical parameters. In all cases the theory is found to recommend a smaller population in stationary state than Total Utilitarianism. The standard of living is consequently higher. Because I make use of global estimates, I am able to lay bare the quantitative differences between the two ethical theories.

Like Total Utilitarianism, Generation-Relative Utilitarianism offers a way to get a measure of the loss that is suffered when someone dies. And like Total Utilitarianism, the theory can be used to get a measure of the loss that would be entailed if the human race were to become extinct. Loss in the latter case would be forward looking; it would reflect the value of lives that would be foregone on account of extinction. In a remarkable work the late Jonathan Schell drew our attention away from that line of reasoning and spoke of the loss each of us would suffer if

[12] In the Appendix the model is extended to include production.

we were to learn that no one will follow us. In Section 11 I review Schell's argument and modify it by enlarging the scope of that loss. By extending the sphere of human motivations we avoid a free-rider problem to which Schell's ethics could be vulnerable. The deep human need to live through time rather than in time is then used in Section 11 to arrive at a view of personal - and thereby collective - stewardship of Earth. Acknowledgement of stewardship points also to an implicit understanding among the generations to protect and promote the biosphere's ability to support life. But it arrives at it from a different direction from the one we are led to by Generation-Relative Utilitarianism. The amended account of personal motivation is not meant to be a substitute for Generation-Relative Utilitarianism, it will be found to be a complement to it. Nor is the implicit understanding among generations contrary to Total Utilitarianism's idea of the universal good. Or so I will argue.

4 The Genesis Problem

Total Utilitarianism, in the form Sidgwick cast it, was applied to a timeless economy by Meade (1955) and extended by Dasgupta (1969) to a world facing an indefinite future. In the latter publication it was shown that the theory encourages large populations. I reproduce that finding by working with a stripped-down version of the model in Dasgupta (1969).

We imagine a timeless world, endowed with a finite stock of resources (Nature) of size K. Because I want the model to reflect the constraints our finite Earth imposes on us over the long run in the simplest possible way, I assume in the text that is no production. The resource serves only as a general consumption good (ecological services). As the world is timeless, stocks (Nature) and flows (ecological services) are the same. In footnote 16 and the Appendix I extend the model to include production.

People are identical. If C is someone's consumption level (C defines her living standard), her personal well-being is $U(C)$. The well-being function U is taken to be increasing and continuous in C, and to increase with C at a diminishing rate.[13] Positive well-being records life

[13] Edgeworth (1881) routinized the assumption of diminishing marginal well-being.

The presumption that the sole factor in well-being is consumption may seem otiose. People have argued that what shapes one's life are personal and social engagements. (There are other factors such as genetic endowment, which at least for now are not subject to choice by the person. We may regard them as parameters of the U-function.) But because all engagements require goods and services, we may think of engagements as production activities, in which goods and services are inputs. Partaking in a meal (alone or in the company of others) is an engagement, but food items are the inputs that make the meal possible. Friendship involves investment in time, which is a scarce resource. And so on. Goods and services come with various characteristics, and for each person there is (given the resources at her command) a best way of obtaining the commodities that are best suited for the activities that are from her point of view the best for her to engage in. Of course, the best way of obtaining goods and services with those characteristics involves further engagements (getting a job, making contacts), and so on; but the regress is circular, meaning that it is closed. The economist's presumption that the only things people care about are goods and services should be seen as a filtered expression of their projects and aims. Commodities have instrumental value; they don't necessarily have intrinsic worth. However, irrespective of what a person values, she will be found to value commodities. The U-

as good, negative well-being records life as bad. U is positive at large values of C but negative at small values of C. It follows that there is a unique value of C at which U is zero. The model being stark, I identify the living standard with consumption and write the standard of living at which $U = 0$ as C^S. Thus $U(C^S) = 0$. Sidgwick would have called C^S the living standard at the "hedonistic zero"; Meade (1955) called it "welfare subsistence". Because the notion of personal well-being here differs from the one adopted by Sidgwick and Meade, C^S is more accurately called "well-being subsistence". Figure 1 depicts a well-being function $U(C)$ with those features.

4.1 Zero Well-Being

Sidgwick (1907: 124-125) spoke of "neutral feeling" when suggesting ways to identify the hedonistic zero:

"If pleasure ... can be arranged in a scale, as greater or less in some finite degree, we are led to the assumption of a hedonistic zero, or perfectly neutral feeling, as a point from which the positive quantity of pleasures may be measured... For pain must be reckoned as the negative quantity of pleasure, to be balanced against and subtracted from the positive in estimating happiness on the whole; we must therefore conceive, as at least ideally possible, a point of transition in consciousness at which we pass from the positive to the negative. It is not absolutely necessary to assume that this strictly indifferent or neutral feeling ever occurs. Still experience seems to show that a state at any rate very nearly approximating to it is even common: and we certainly experience continual transitions from pleasure to pain and *vice versa*, and thus (unless we conceive all such transitions to be abrupt) we must exist at least momentarily in this neutral state."

Zero well-being is a defining notion in population ethics. Sidgwick's reference to neutral feeling invites us to assess life from the inside. But the idea of neutral feeling also points, at least tangentially, to a comparison of life with non-existence. That latter exercise requires calibrating well-being in terms of something outside our experience. Nagel (1979) famously suggested that death is not an unimaginable condition of the living person, but a mere blank, and that it can have no value whatever, positive or negative. I can't tell whether by that Nagel meant the blank can't be assigned the number zero, as a benchmark against which other states of affairs are compared, but he went on to suggest that one can imagine non-existence by imagining being in a coma for the rest of one's life. For Nagel non-existence is the real blank, being totally unconscious for the rest of one's life is a simulation of that blank. Even that simulated state takes some strain to

---

function in the text should be interpreted in that light. I make the innocuous further assumption that $U(C)$ is a differentiable function of C. That enables us to use the calculus for solving for the optimum population size.

imagine, but it isn't absurd to imagine it.[14]

We are thinking of someone's life as a whole, not the quality of her life at a moment in time. There are thinkers who believe the whole is the sum of its momentary parts; there are others who believe the whole to be a non-linear functional of those parts. We don't need to adopt a position on that because excepting for the model we analyse in Sections 8 and 10, there will be no occasion to slice someone's life into parts. Lives are judged to be good or bad in comparison to non-existence. That's an essential feature of population ethics, because the grammar we construct has to offer a criterion with which to judge when creating a person is justified. The U-function here is calibrated by using non-existence as a point of reference. Without loss of generality we attach the number 0 to that point. Zero well-being is therefore the measure of a life that, taken as a whole, goes neither well nor badly. In view of the additive structure that characterises Total Utilitarianism's conception of aggregate utility, $U = 0$ is also the level of well-being at which an additional life adds no additional value to the population that contains it.

In conversations I have heard it suggested that $U = 0$ is the point of indifference between dying and continuing to live, or the point of indifference between life and death. In Section 5 we uncover the reason that interpretation is entirely misconceived. That reason also steers us away from the thought that C^S is a poverty line, in the sense the notion has been articulated in such figures as the World Bank's "1.25 dollars-a-day", or even the "two-dollars a day" benchmark. Every life of poverty represents a bad state of affairs for that individual and makes the entire population less valuable. C^S should be sought in a higher range of figures, possibly a lot higher than the World Bank's poverty line. C^S is the living standard at which life is neither good nor bad. No doubt it is difficult to determine it, but the idea underlying well-being subsistence cannot be bypassed in population ethics. Identifying C^S involves a deep and difficult value judgment.

4.2 The Sidgwick-Meade Rule

In applying Total Utilitarianism's reasoning to population ethics it is simplest to apply it first to the Genesis Problem. Genesis is also the scene where Total Utilitarianism doesn't run into any meaningful trouble. So we imagine Earth to be devoid of people.

Let N be the number to be created. Because marginal well-being declines as the living standard increases, an equal distribution of K among all who are created is the ideal. If N people were created each would receive K/N units of the consumption good. Thus $C = K/N$. Total Utilitarianism tells us to search for the value of N at which $NU(K/N)$ attains its maximum value.

Routine arguments show that if K is large (N in that case can be regarded as a continuous variable), optimum N satisfies the condition,

---

[14] Nagel (1979: 2): "... the value of life and its contents does not attach to mere organic survival: almost everyone would be indifferent (other things equal) between immediate death and immediate coma followed by death twenty years later without reawakening." I develop this point in Section 5, but apply it to a different purpose from Nagel's.

$$dU(C)/dC = U(C)/C. \quad (1)$$

Equation (1) says that the solution of the Genesis Problem is a population size at which marginal well-being equals the ratio of average well-being to average consumption (Figure 1). The equation is fundamental to Total Utilitarianism. We will call it the Sidgwick-Meade Rule.

The intuition behind the Rule is this:

At the Genesis Ideal neither a small hypothetical increase in population nor a small hypothetical decrease would alter total well-being. Suppose now a marginal increase is contemplated. (The argument associated with a marginal decrease is analogous.) The additional person would share *K* equally with the population that was originally contemplated. The increase in total well-being that would obtain if that additional person was created is her well-being, which is $U(C)$. But there would also be a decrease in total well-being, which is that each of the others would have slightly less consumption. That potential loss in well-being is $C[dU(C)/dC]$. At the Genesis Optimum the potential gain and the corresponding potential loss in total well-being must be equal. The Sidgwick-Meade Rule asserts this. We denote the solution of equation (1) by C^O and the corresponding population size by N^O. It follows that $N^O = K/C^O$. N^O is the ideal population size and C^O is the ideal living standard.[15]

4.3 Measurability, Comparability, and the Aggregation of Personal Well-Beings

There is no absolute scale for measuring personal well-being. The theory of choice under uncertainty, for example, yields a measure that is unique up to positive affine transformations. That says if U_i is a measure of individual *i*'s well-being and if α is a positive number and β is a number of either sign, then $\alpha U_i+\beta$ is an equally valid measure of *i*'s well-being. Put another way, the ordering of states of affairs represented by U_i is the same as the ordering represented by $\alpha U_i+\beta$. In the latter scale α represents the unit and β the level. We say U_i in that case is a "cardinal" measure. Concerned as it is with well-being differences among alternative states of affairs, the theory does not require the notion of zero well-being.[16]

---

[15] Dasgupta (1969) assumed that reproducible capital, a fixed quantity of land, and labour produce an all-purpose good that can be either consumed or accumulated as (reproducible) capital. To see what the Sidgwick-Meade Rule looks like when there are production possibilities, consider a simple version of that model. Suppose output *Y* is a function of the biosphere (*K*) and population (*N*), so that we may write $Y = F(K,N)$, where *F* is an increasing function of *K* and *N*. *K* is fixed by assumption. In the timeless economy $C = F(K,N)/N$. It is now easy to confirm that the Sidgwick-Meade Rule is

$$dU(C)/dC = U(C)/[C - \partial F/\partial N] > U(C)/C. \quad (F1)$$

Compare equation (F1) with equation (1) in the text. The Rule now says that marginal well-being of consumption exceeds average well-being per unit of consumption. Other things equal the Total Utilitarian optimum recommends a larger population and a lower living standard in a world where people are not merely consumers but are producers too, than in a world where they are mere consumers.

[16] In his classic work, Harsanyi (1955) built Utilitarianism on the basis of the theory of choice under uncertainty.

Population ethics in contrast requires that lives that go well be distinguished from those that go badly. A life that goes neither well nor badly scores zero (in Sidgwick's system of ethics it is the hedonistic zero). Personal well-being is "strongly cardinal" if its measure is unique up to positive linear transformations (i.e. proportional transformations). That says if U_i is a measure of *i*'s well-being and if α is a positive number, then αU_i is an equally valid measure of *i*'s well-being. A population ethic that is derived from Total Utilitarianism requires personal well-being to be strongly cardinal.

When applying Total Utilitarianism to the Genesis Problem we made an implicit assumption, that well-beings are fully comparable across individuals. There was no need to mention it because people were assumed to be identical. In a heterogeneous population the assumption should be made explicit. Without full comparability of personal well-beings it would not be possible to construct an ethics with which to derive population policies.[17]

The idea of full comparability is familiar from weights and measures. The weights of objects (measured, say, in a vacuum-sealed flask at ground level in a given latitude) are fully comparable. Suppose we find *x* to be a heavier object than *y*. If that is to be a meaningful finding, the units in which they are measured must be the same; it's no good measuring *x* in ounces and *y* in grams. Reference to grams and ounces tells us that we can say a lot more than merely that *x* is heavier than *y*: we can say how much heavier *x* is proportionately than *y*. The reason we can is that if *x* is found to be twice as heavy as *y* using one system of units (ounces), it will be found to be twice as heavy as *y* using any other system of units (grams). And that's because an ounce is proportional to a gram. We can move from one system of units to another with impunity so long as the corresponding transformations (grams to ounces) are applied consistently.

We can say even more. Because physical theories tell us that addition and subtraction are legitimate (even required) operations on weights, each scale in the set of admissible scales is proportional to any other scale in the set, and all scales that are proportional to a scale in the admissible set are also in the admissible set. Non-linear scales don't belong.

We say personal well-beings are fully comparable if multiplying each individual's well-being by a constant positive number α is ethically of no significance. Other things equal the Total Utilitarian would rank the pair of personal well-beings {4,11} above the pair {5,9}, because 15 is bigger than 14. Full comparability says that if we multiply each individual's well-being by a positive number α, the ranking of the resulting pairs remains the same. They remain the same because 4α+11α (= 15α) is bigger than 5α+9α (= 14α) for all positive α. The Total Utilitarian can choose any value of α she likes without compromising her ethics.

---

[17] Sen (1970) is the classic on the measurability and interpersonal comparability of individual well-beings in fixed populations. By extending his analysis to population ethics, it is easy to show that dropping full comparability would yield only a partial ordering of population policies. Partial orderings may be the best that the ethicist can hope for in practice, but no theory of social ethics should start with partial orderings.

Total Utilitarianism aggregates personal well-beings by summing them. The theory requires personal well-beings to be strongly cardinal and interpersonally fully comparable.[18] The reasoning confirms that the Sidgwick-Meade Rule is invariant with respect to α.[19] Notice moreover that well-being subsistence, C^S, remains the same under all proportional transformations of the *U*-function. That's because $\alpha U(C^S) = 0$ for all positive α. In the model we used here to study the Genesis Problem, α was set equal to 1 so as not to add another symbol.

4.4 Large Populations Under Total Utilitarianism

Aggregate well-being in Total Utilitarianism is the product of population size (*N*) and average well-being (*U*). Rawls (1972: 162-163) noted that, as shown in Figure 2, the tradeoff between *N* and *U* in the product *NU* is such that no matter how small is *U*, so long as it is positive, a sufficiently large *N* can compensate for a reduction in *U* (as long as it remains positive of course). Parfit (1984: 425-441) found that feature of Total Utilitarianism repugnant. So he called it the Repugnant Conclusion (RC). Previously Dasgupta (1969) had noted that Total Utilitarianism commends large populations, in the sense that it advocates population sizes for which the (optimum) standard of living, C^O, is only a little above well-being subsistence, C^S. The analysis also showed that the larger is the elasticity of marginal well-being with respect to personal consumption, the lower is the optimum living standard, and that the latter tends in the limit to the standard of living at which well-being is zero. To reconstruct the argument in the timeless economy, suppose as in Dasgupta (1969) that

$$U(C) = B - C^{-\sigma}, \qquad B > 0 \text{ and } \sigma > 0. \tag{2}$$

In equation (2) $U(C)$ is defined by two parameters, *B* and σ, both of which are positive numbers (Figure 3). *U* is an increasing function of *C* and is bounded above by *B*. $1+\sigma$ (> 1) is the elasticity of marginal well-being with respect to consumption. Ramsey (1928) named *B*, Bliss.[20]

Routine calculations show that on using equation (2) in the Sidgwick-Meade Rule,

$$C^O/C^S = (1 + \sigma)^{1/\sigma} > 1. \tag{3}$$

---

[18] The reverse, much deeper question - on the limits to the forms of aggregation that are permitted of personal well-beings when the latter are strongly cardinal and are fully comparable - was posed and answered in a remarkable work by Blackorby and Donaldson (1982). The showed that summing personal well-beings is certainly an admissible way to aggregate, but that there are other forms of aggregation (e.g., summing weighted well-beings).

[19] If $U(C)$ is replaced by $\alpha U(C)$ in equation (1) and $dU(C)/dC$ by $d[\alpha U(C)]/dC$, α cancels from the two sides.

[20] The *U*-function in equation (2) is unbounded below. In applied work that should be an unacceptable feature of the *U*-function because unboundedness is inconsistent with the theory of decision under uncertainty (Arrow, 1965). But here we are engaged in theoretical reasoning. The *U*-function in equation (2) has the virtue of revealing the structure of Total Utilitarianism in a very simple manner.

Equation (3) is informative because it relates the optimum living standard to well-being subsistence.[21] For population ecologists however, it proves more natural to recast the equation in terms of population size. With that in mind define $N^S = K/C^S$. Because I am thinking of the global economy here, N^S can be interpreted as the biosphere's human *carrying capacity*.[22] Equation (3) can then be re-expressed as

$$N^S/N^O = (1 + \sigma)^{1/\sigma} > 1. \quad (4)$$

Recall now that if σ is positive, $(1+\sigma)^{1/\sigma}$ is less than e (the base of natural logarithms), which in turn is approximately 2.74. It follows from equations (3)-(4) that

$$C^O/C^S = N^S/N^O = (1 + \sigma)^{1/\sigma} < e \approx 2.74. \quad (5)$$

Equation (5) could be taken to say that Total Utilitarianism favours large populations. No matter how large or small σ happens to be, the optimum living standard is less than 2.74 times well-being subsistence. The equation also says that carrying capacity is less than 2.74 times the optimum population size. To illustrate, suppose $\sigma = 1$. Then $C^O/C^S = N^S/N^O = 2$, which means the optimum living standard (C^O) is twice well-being subsistence, and carrying capacity is twice the optimum population size (N^O). If the biosphere's carrying capacity is reckoned to be 10 billion people, the optimum population size would be 5 billion.

To be sure, the idea of Earth's carrying capacity points to a range, not an exact figure. Technological advances and improved efficiency in the way resources are used raise the biosphere's carrying capacity. Moreover, the notion is not independent of ethical values: N^S is an inverse function of C^S. The point however is that Earth's life support system is bounded, and N^S captures that latter fact. That said, whether Total Utilitarianism should be seen to be advocating overly large populations ("pro-natalist", as some would say) remains unclear. Much hangs on the standard of living C^S. Identifying C^S involves a value judgement. Once it has been chosen, the factor that remains to determine C^O is σ. That means K plays no role in determining the ideal standard of living. If K was to be larger, Total Utilitarianism would respond by adjusting the recommended population size N^O.[23] This could certainly be seen as an expression of pro-natalism;

---

[21] To confirm equation (3), notice first that $C^S = B^{-1/\sigma}$. Next, use equation (2) in the Sidgwick-Meade Rule to show that $C^O = B^{-1/\sigma}(1+\sigma)^{1/\sigma}$. Equation (3) follows immediately.

[22] Ehrlich and Holdren (1971) is the classic that introduced the metaphor, I=PAT, to draw attention to the significance of Earth's carrying capacity for population ethics. The authors traced the *i*mpact of human activities on the Earth system to *p*opulation, *a*ffluence (read, the standard of living), and the character of *t*echnology in use (including knowledge, institutions, social capital). Contemporary demographers have in large measure ignored the Ehrlich-Holdren insight that nature doesn't calculate percentages (e.g. declines in the rate of growth of population), she responds to absolute quantities. The notion of Earth's (human) carrying capacity is further developed in Section 9, where I use international statistics to put flesh into population ethics.

[23] I owe this observation to Kenneth Arrow.

but pointing in the opposite direction gives an entirely different impression. Suppose K was to be smaller. The same feature of equation (2) now says that the adjustment to the ideal would be made entirely through a reduction in population size; the standard of living would be fully protected. When applying the theory to global economic statistics in Section 10, I return to this feature of the calculus of Total Utilitarianism.

$1+\sigma$ is the percentage rate at which marginal well-being declines in response to a percentage increase in the living standard. Value pluralists have used $1+\sigma$ also as a measure of the aversion to inequality in the distribution of living standards. When personal well-being is identified from attitudes to risk in the standard of living, $1+\sigma$ is also a measure of a person's risk aversion.[24] To obtain a feel for the way $1+\sigma$ influences the optimum living standard, we subject the Sidgwick-Meade Rule to a full sensitivity analysis. To do that we compute C^O/C^S in equation (5) using alternative values of σ. It will be noticed that to vary σ is to vary well-being subsistence (equation (2)), but that's an acknowledgement that determining well-being subsistence involves a value judgment.

How does C^O vary in relation to C^S as σ assumes different values? Equation (5) says that the larger is σ, the closer is C^O/C^S ($= N^S/N^O$) to 1.[25] If carrying capacity is 10 billion people and σ is very large, the optimum population size is close to 10 billion. For the value pluralist, strong social aversion to *consumption* inequality recommends very large populations. I have known this result since my student days but still find it puzzling that inequality aversion (respectively, risk aversion) should play so influential a role in applied population ethics. Admittedly the theory we are invoking here says that all who are born are to be treated equally; even so, it isn't *a priori* obvious why an attitude toward consumption inequality (alternatively, consumption risk) should so influence the ideal size of population. (*Ex post* it is obvious of course.)

I have never understood why the large populations that are advocated by Total Utilitarianism should be found repugnant. So long as well-being is positive, a person's life is good. Life may not be very good, but it is good. The tradeoff between population size and average well-being in Total Utilitarianism may not be to everyone's liking, but it is all too precious to be repelled by it.[26] The issue isn't a matter of semantics, nor does it hang on any

---

[24] Vital contributions to the modern literature on income inequality measures are Kolm (1969) and Atkinson (1970) and to the measurement of risk aversion are Pratt (1964) and Arrow (1965). Dasgupta (2008) offered reasons why the plausible range of values for σ is (0,2]. Empirical studies of choice under uncertainty have usually revealed σ to lie in the range (0,1].

[25] That's because $(1+\sigma)^{1/\sigma} \to 1$ as $\sigma \to \infty$.

[26] The Repugnant Conclusion initiated a large literature. See for example McMahan (1981), Sikora (1981), Hurka (1983), Sterba (1987), Temkin (1987), and Cowen (1989). Huemer (2018) contains a similar reaction to the Repugnant Conclusion to mine.

There is an economics literature from the first half of the twentieth century that took J.S. Mill's Average Utilitarianism to be the basis of population ethics (Gottlieb, 1945, is a well-

particular notion of well-being. Once the ethicist settles on a figure for well-being subsistence, she simply has to acknowledge that life is good at any standard of living that exceeds it (the higher, the better, of course). It is time to retire the Repugnant Conclusion from population ethics.[27]

We now study a possible reason Parfit and ethicists following him have found the trade-off in Sidgwick's Utilitarianism between population size and average well-being to be repugnant. We will confirm that the reason is based on a misreading of zero well-being.

5 Death

In a moving discourse on the place of autonomy and responsibility in personal well-being, Williams (1993: 50-102) drew attention to an aspect of personal responsibility that starts not from what others may demand of someone, but from what that someone demands of himself. Williams reminded readers that Sophocles had reported that Ajax, being slighted by the award of Achilles' arms to Odysseus, had intended to kill the leaders of the Greek army. To prevent the massacre, Athena made Ajax mad. It is significant that Ajax's condition didn't affect his purposes; rather, it altered his perception. Thinking that he was killing Odysseus and the others, Ajax slaughtered the army's flock of sheep and cattle. In Sophocles' account the despair arising

---

known account of the literature). But Average Utilitarianism faces insuperable difficulties. In place of the Sidgwick-Meade Rule (equation (1)), Average Utilitarianism would recommend a population of size 1 in our model.

[27] Blackorby and Donaldson (1984) proposed that other things equal the creation of an additional person should be judged to be good only if her well-being were to exceed a positive, critical level U^C (> 0). The authors called their theory "Critical-Level Utilitarianism". (See also Blackorby, Bossert, and Donaldson, 1997; and Broome, 2004 (pp. 199-202) for a supportive account of the theory). U^C is not identified from any moral reasoning; its sole purpose would seem to be to offer an escape from the ubiquitous "Repugnant Conclusion". But there is a basic problem with the theory. Suppose the well-being of each existing person is expected to be $2U^C/3$. As this is a positive number, everyone alive expects a good life. Critical-Level Utilitarianism asks us to believe that other things equal creating an additional person whose well-being would be $2U^C/3$ is a bad idea. The authors don't explain why it would be a bad idea. I don't see why it wouldn't be a good idea. Ng (1986, 1989) is one of the few writers to have insisted that if U is positive, life is good.

Ingmar Schumacher has suggested in correspondence a possible re-interpretation of the Blackorby-Donaldson theory. We are to imagine that the authors re-calibrated the personal well-being function U, so that their U^C is the hedonistic zero (i.e. Classical Utilitarianism's $U = 0$). In that interpretation life is bad if U lies in the interval $[0,U^C)$ and $U = 0$ reflects a condition where a person contemplates putting an end to his life. But if Schumacher is right, Critical-Level Utilitarianism is simply a renumbering of Total Utilitarianism, and that would imply it cannot act as an antidote to the Repugnant Conclusion.

from the shame Ajax felt on awakening left no option open to him but to take his own life. And Williams observed (p. 76) that when Ajax says he must go, "... he means that he must go: period."

Modern ethicists in contrast have frequently offered a view of life that is not only at odds with Sophocles' account but also with commonly-held notions of life's quality and of the reasons for the fear and horror that death holds for us. In the chapter that introduces Utilitarianism to his readers, Sidgwick (1907: 414-415) wrote:

"... I shall assume that, for human beings generally, life on the average yields a positive balance of pleasure over pain. This has been denied by thoughtful persons: but the denial seems to me clearly opposed to the common experience of mankind, as expressed in their commonly accepted principles of action. The great majority of men, in the great majority of conditions under which human life is lived, certainly act as if death were one of the worst of evils, for themselves and for those whom they love..."

Nagel (1979) concluded that if death is an evil, it is the loss of life that is objectionable. The conclusion is incontrovertible to the secular mind, but there are at least three circumstances of death that should be distinguished, and they don't point in the same direction. There is death that comes naturally to one in the fullness of time; there is death that comes from natural causes before one's time; and there is death that is brought on one by one's own deliberate action. Nagel contrasted the first two, but didn't speak to the third. And it is the latter that should make us pause before accepting Sidgwick's conclusion that life, all in all, is a positive good for most people.

Religious prohibition, fear of the process of dying (the possibility of suffering pain, the feeling of isolation), the thought that one would be betraying family and friends, and the deep resistance to the idea of taking one's own life that has been built into us through selection pressure would cause someone even in deep misery to balk. It may even be that no matter what life throws at us we adjust to it, if only to make it possible to carry on.[28] But the acid test for the conclusion that life, all in all, is invariably a good is to ask ourselves whether we shouldn't pause before creating a person so as to imagine the kind of life that is likely to be in store for the potential child.[29] The desire to procreate springs from our deep emotional needs, and the direct motivation we have to create children can be traced to a wide variety of reasons (we noted a few in Section 1), but here we are concerned only with the life of the prospective child.

---

[28] Suicide isn't that rare a phenomenon. Annually over 1 million people take their own life worldwide (approximately 1.8 per cent of deaths are suicides) and 20-30 million people attempt suicide.

[29] Nagel acknowledges that the loss encountered when someone's life is cut short can't be balanced on a one-to-one basis by replacing that someone with a new birth. In my readings on the subject I have found him (and Narveson, 1967, and Heyd, 1992) to be among the very few who (at least until recently) acknowledged the asymmetry. Below (Section 6) I make use of the asymmetry in developing population ethics.

The view that life overall is always good have led philosophers to conclusions that should be unacceptable. In the passage in which he presented the Repugnant Conclusion, Parfit (1984: 388) recognized Total Utilitarianism's pro-natalist character, but then interpreted it in the following way:

"For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living."

One wouldn't question that the image is repugnant, but it belongs nowhere, certainly not in Total Utilitarianism. We are asked to consider a figure for world population that will almost certainly be reached in the second half of this century (a figure that is most unlikely to be sustainable at reasonable material comfort - Section 10), and are then made to imagine an Earth where, because of population pressure, people scramble for resources so as to eke out an existence, having *lives that are barely worth living*.

But someone whose life is barely worth living doesn't enjoy a life of positive quality, she suffers from *negative* well-being. In the contemporary world over a half billion people are malnourished and prone regularly to illness and disease, many of whom are also debt ridden, but who survive and tenaciously display that their lives are worth living by the fact that they persist in wishing to live. If you were to say that you would not wish the circumstances they endure on anyone, I wouldn't take you to mean their lives aren't worth living; I would take you to be saying that their circumstances are so bad that you wouldn't wish them on even your worst enemy, that something ought to be done to improve their lives, that if you were to disregard the countervailing needs you and your household may have, you wouldn't want to *create* children facing those circumstances.[30]

Death *relieved* the intolerable pain Ajax experienced on awakening from the madness Athena had inflicted on him. Ajax knew it would, which is why he chose it. It was better for him that he paid the price of death than that he carried on. The inference Sidgwick drew from the fact that death is generally thought to be one of the worst evils is altogether unfounded, and I cannot imagine how so profound and careful a thinker could have made such an elementary arithmetical error. That death is a horror to most people doesn't imply that life is on balance pleasurable. On the contrary, the greater is the horror that taking one's own life poses to someone (betrayal of one's family and friends, revelation of one's misery to others when one wants it to remain undisclosed even after death), the *more* he would be willing to carry on in a state of misery. One way to interpret life in the range between the point at which a person takes his life and well-being

[30] Parfit's interpretation of well-being subsistence wouldn't appear to have changed over the years. On the Repugnant Conclusion, he has written recently, "Compared with the existence of many people who would all have some very high quality of life, there is some much larger number of people whose existence would be better, even though these people would all have lives that are barely worth living." Parfit (2016: 110).

subsistence is to view it as bearable but not good. The person would not contemplate suicide, but could entertain the feeling that it would have been better not to have been born. Zero well-being is then the transition point from the bearable to the good.[31] That is why estimating well-being subsistence from people's behaviour or responses to questionnaires would be a mistake.[32]

6 The Problem of Sleeping Beauty

In the Genesis Problem Earth is devoid of people. The domain of discourse consists only of potential persons. In any Actual Problem Earth is inhabited. The question is whether further people should be created. The two settings are wildly different.

Consider the problem of Sleeping Beauty. She is alive but in a state of total unconsciousness (Sidgwick would regard her state to be the hedonistic zero). What makes her an interesting case for us here is that her life expectation is the prevailing average, in that with a modicum of medical attention she can be expected to complete the natural lifespan in a state of total unconsciousness. A small expenditure can however revive her immediately and fully, in

---

[31] I owe this interpretation to Robert Solow.

[32] To illustrate Sidgwick's error, imagine that in the units chosen to measure U, the horror of suicide for someone is -300. The person would choose to continue to live so long life offered her a value in excess of -300; and that could be as low as -299.99. The presumption that life on balance is inevitably good ($U > 0$) or that it is invariably bad ($U < 0$) precludes reasoned discussion on population policies. The former commends a large (even infinite) population; the latter commends an Earth devoid of people. That alone shows that unlike decision theory, population ethics requires of us to specify the zero of the U-function.

As far as I can tell, numerical measures of subjective well-being have also been taken to be non-negative in all the large-scale surveys on happiness and life satisfaction that have been conducted in recent years (see Helliwell, Layard, and Sachs, 2013, for an account of the findings). It would look as though designers of the questionnaires that have shaped the literature on empirical hedonism have also been much influenced by the thought that life satisfaction can never be negative (otherwise why do people continue to want to live?). We do not know whether people would respond differently to questionnaires on life satisfaction if, instead of a scale from, say 0 to 10, respondents were offered a scale from -5 to +5.

At the other extreme from Sidgwick is the Chorus in *Oedipus at Colonus*, who were reported by Sophocles as having observed: "Not to be born is, past all prizing, best; but, when a man hath seen the light, this is next best by far, that with all speed he should go thither, whence he has come." (I am grateful to Kenneth Arrow for this reference. The translation is R.C. Jebb's.)

Daly et al. (2011) have found in one study that a major determinant of suicide is being at a well-being level significantly lower than others, thus implying that living in a country where well-being is high is a risk factor for committing suicide. This can be explained if the coefficient of variation in incomes in two countries, one rich the other poor, is the same and personal well-being accords with Veblen (1924), that, other things equal U is a decreasing function of the average living standard relative to a person's own. See Arrow and Dasgupta (2008) for an analysis of the optimum rate of saving in a world where U is an increasing function not only of one's own living standard but also her living standard relative to the average living standard of others in her society. Veblen's observation on human psychology found a telling expression in a remark attributed to a Garry Feldman of Stamford, Connecticut, one of the wealthiest towns in the USA: "I might be in the top one per cent, but I feel that I am in the bottom third of the people I know." (*The Guardian*, Saturday 16 February 2013).

which case it is confidently expected her lifetime well-being will be U^* (> 0). That's option X for her parents. Option Y is for them to conceive another child who would enjoy a lifetime well-being equal to U^*. However, under Y Sleeping Beauty will remain unconscious. We suppose that the couple have no special feelings for Sleeping Beauty. She was born unconscious, so they have never got to know her. Assume now that in all other respects X and Y have the same consequences. What should Sleeping Beauty's parents do?

If, as in a literal interpretation of Sidgwick's Utilitarianism, agreeable consciousness is the sole good and if the fact that something good would be the result of one's action is the basic reason for doing anything (the ground of binding reason) - the two together yield the principle of universalizability in ethical reasoning - the couple in question should be indifferent between X and Y. But there are a number of reasons X should be viewed as the right option. They all involve constraining the consequentialist reasoning I have been following here with deontological directives.

6.1 Imperatives

I first list three reasons that have been deployed by philosophers studying problems similar to that of Sleeping Beauty. I then work with a fourth reason (Section 7), based on Arrow (1981) and Scheffler (1994). There I borrow the Agent-Based perspective Scheffler constructed and apply it to population ethics.

Parental Obligation to Sleeping Beauty

X is the right option because Sleeping Beauty's parents have an obligation toward her which they don't have toward a potential child. They have an obligation toward Sleeping Beauty because they were responsible for conceiving her. People don't have an obligation to become parents, but they acquire one if they choose to become parents. It has been argued that parents have an obligation toward their offspring that no one else has. Parental obligation provides an argument for choosing X over Y.[33]

Sleeping Beauty's Claim

Sleeping Beauty has claims over her parents. She is a person, it is *her* well-being that is subject to parental choice. She has a claim to be acknowledged by her parents; she has a claim to be heard by them even though she is unable to speak. The potential child has no comparable claims.[34]

---

[33] In his comments on the population problem Seabright (1994) framed parental obligation in terms of the parents' notion of "family". He argued that the family has a special place in ethics.

Simon Beard has pointed out to me that the example can be expanded to encompass a wider range of obligations, for example, if the chooser between X and Y was Sleeping Beauty's friends (they would have no obligation to make a new friend).

[34] Gosseries (2010) has a good discussion of this. He studies the obligations people have to one another, not simply parents toward their children. I am ignoring the obligation people may have (at least to themselves) to procreate in special cases, such as a population at risk of

The Value of Sleeping Beauty's Life

Nagel (1979:2) suggested that life has a value to the person living it that is independent of anything she may experience:

"The situation is roughly this: There are elements which, if added to one's experience, make life better; there are other elements which, if added to one's experience, make life worse. But what remains when these are set aside is not merely *neutral*: it is emphatically positive. Therefore life is worth living even when the bad elements of experience are plentyful, and the good ones too meager to outweigh the bad ones on their own. The additional positive weight is supplied by experience itself, rather than by any of its contents."

I am not sure I understand what Nagel meant, but one interpretation is that life has a value independent of whether it goes well or badly (it is "supplied by experience itself, rather than by any of its contents").[35] Call that stand-alone value, Ω. It sits outside the calculus of well-being. When we say life is precious, even sacred, we point to Ω. That life has a stand-alone value Ω is not a reason for creating a person, but is a reason for awakening Sleeping Beauty even if her life were predicted not to be one of positive quality.[36] Ω enters the moral calculus only when it is applied to actual persons. Sleeping Beauty is a person, in contrast to an "unborn person", who is no one. Her parents can, as can others such as the doctors caring for her, point to Ω and require that she be revived.

6.2 Wider Cases

The problem of Sleeping Beauty gives rise to a number of variations. Suppose *X* is not an option. Her parents, or for that matter anyone else, could argue that she should be released from the indignity of living in coma for the rest of her natural life. In that case they would give their consent to having her life-support system disconnected.[37]

Consider now a case where *X* is not available, but *X*' is. In *X*' Sleeping Beauty is revived at a small cost, but her life will not be good. Even though she will be able to function, she will suffer from discomfort and pain and will not lead a fulfilling life. All in all it is expected that her lifetime well-being will be *U***, a not-too-large negative number. It can be argued, and

---

extinction from non-fecundity. See Section 11, on existential risks.

[35] When I first read the passage just quoted, I interpreted it differently. I took Nagel to be saying that (invariant) value is so high that no matter how terrible the person's experiences may be throughout her life, the net value is positive. For the claim to be believable, one must imagine with Nagel (although Nagel left the thought unsaid) that the value of life for a person, qua life, is infinite. To me that read odd. I now believe I have interpreted Nagel correctly and am most grateful to Itai Sher for pointing me in the right direction on this.

[36] Provided of course that her life will not be of utter misery.

[37] It bears emphasis that removing Sleeping Beauty's life support system is not the same as killing her.

contemporary sensibilities would favour the argument, that she be revived, on grounds that Sleeping Beauty should be allowed to exercise her agency, her right to live a life. That's one way to justify *X*'.

Consider yet another variation of the problem of Sleeping Beauty. Option *Y* remains, but in place of *X* we are faced with *X*", which involves reviving Sleeping Beauty with the understanding that her lifetime well-being will be $U^*-\varepsilon$, where ε is a small positive number. Assuming their ethical ordering over states of affair is continuous, her parents (or society, more generally) would choose *X*" if ε was sufficiently small. By continuity, there is a value of ε, say ε^*, such that it is a matter of ethical indifference to them whether *Y* is chosen or whether Sleeping Beauty is revived with the understanding that her lifetime well-being will be $U^*-\varepsilon^*$. To acknowledge that ε^* is positive is to say that, other things equal, reviving Sleeping Beauty with an expectation of $U^*-\varepsilon$ (where $0 < \varepsilon < \varepsilon^*$) would have priority over creating a person with a lifetime well-being of U^*.

And so on. None of these cases is straightforward, but that only shows our intuitions over birth and death are unsteady. One point however is incontrovertible: A person could regard her life to be worth living even if she expects it to be one of misery.

6.3 Agent-Centred Restrictions

Scheffler (1994) uncovered agent-centred restrictions that people can justifiably use when deliberating over courses of action open to them. In a society of *N* persons those restrictions would apply reciprocally, meaning that the state of affairs that would ensue would be the outcome of *N* choices, each having been being guided by an agent-relative perspective.[38]

If Sleeping Beauty was a Total Utilitarian and was able to exercise her agency, she would have faced the dilemma of reconciling the place of the universal others with the demands that her self requires her to meet when faced with morally significant choices. But Sleeping Beauty is unable to exercise her agency. So the choice between *X* and *Y* requires of others, not just her parents, to accommodate the agent-centred restrictions that she could justifiably deploy if she were able to choose. In what follows we study how such restrictions modify Total Utilitarianism.

In the models to follow we are to imagine that each generation reasons collectively. The agent in each generation is therefore that "generation", not the set of all generations beginning with that one. Because generations enter sequentially, the agent-centred restrictions we study are unidirectional, not reciprocal. Moreover, the restrictions are only a weak version of Scheffler's agent-relativism; they only involve the place of potential well-beings in present people's moral sensibilities. When applied to Utilitarianism, those restrictions translate to a weighted sum of

[38] Arrow (1981) used this reasoning to develop a theory of income distributions arising from voluntary contributions. He (personal communication) was moved to construct Agent-Relative Utilitarianism by the words of the first century sage, Rabbi Hillel: "If I am not for myself, then who is for me? And if I am not for others, then who am I? And if not now, when?"

personal well-beings: the weight awarded to potential well-beings is lower than that awarded to present and future people. We will call that perspective Generation-Relative Utilitarianism.[39]

7 Actual Problems and Generation-Relative Ethics

I first put the theory to work on the timeless model of Section 4. I then develop the theory for an economy moving through time.

7.1 Population Ethics in a Timeless World

The world is inhabited by N_0 identical people. The total quantity of consumption available is K. As before, if C is someone's living standard, her well-being is $U(C)$ in the chosen scale. The question is whether creating further people would make a better world. To have an interesting problem, we assume that the world is thinly populated; which is to say, K is large relative to N_0. If they were not to create further people, the well-being of each of the N_0 persons would be high.

Suppose an additional N_1 people are considered. Once they are born they assume the status of present people. They would then receive the same consideration as the N_0 existing persons.[40] Straightforward Utilitarian arguments imply that if they were created, the additional N_1 people would share K equally with the N_0 people who created them. Each person's living standard would then be $K/(N_0+N_1)$. The backward-induction I have just deployed is essential to Generation-Relative Utilitarianism.

We assume that the weight awarded to the well-being of present people is taken to be 1 and the weight awarded to the well-being of potential people is π, where π is a number between 0 and 1. Denote aggregate well-being as judged by the existing N_0 people by V. Then

$$V = N_0U(K/(N_0+N_1)) + \pi N_1U(K/(N_0+N_1)). \quad (6)$$

What we may call an Actual Problem (as opposed to the Genesis Problem) is to choose N_1 so as to maximize V in equation (6).[41] Routine calculations show that under Generation-Relative Utilitarianism the reproductive rule that corresponds to the Sidgwick-Meade Rule is

---

[39] The considerations in Sections 6.1-6.2 don't imply Generation-Relative Utilitarianism, but they come close to it. They are also a reminder that the pure consequentialist thinking underlying Total Utilitarianism requires deontological directives if we are to make sense of ethical dilemmas in extreme settings, like birth and death. I am grateful to Itai Sher for correspondence on this.

[40] Obviously, if K/N_0 was small, the optimum value of N_1 would be 0.

[41] π corresponds to ε^* in Section 6.2. Notice that the larger is ε^*, the smaller is π.

V in equation (6) is a direct extension of Total Utilitarianism and would reduce to it if π were to be set equal to 1. I have arrived at it by applying Scheffler's reasoning to population ethics. Value Pluralism in contrast would require the formula to be derived from ethical axioms on the way individual U-functions are combined to yield a social well-being function V. Welfare axiomatics does that. Blackorby and Donaldson (1984) applied welfare axiomatics to population ethics. Their techniques can be adapted to yield equation (6).

$$dU(C)/dC = [\pi(N_0+N_1)/(N_0+\pi N_1)][U(C)/C] < U(C)/C. \quad (7)$$

Equations (1) and (7) say that the optimum standard of living in the Actual Problem (label it C^{OO}) is larger than in the Genesis Problem (C^{O}). That is shown in Figure 4. It follows that the optimum population size is smaller in an Actual Problem than in the Genesis Problem. The smaller is π, the fewer are the additional numbers (N_1) that would be recommended by Generation-Relative Utilitarianism.

We may put the matter in a slightly more technical, way: There exists a function $N(\pi)$, which decreases with π, such that if $N_0 < N(\pi)$, equation (7) yields the optimum population size. But if $N_0 > N(\pi)$, the optimum policy is to not add to the existing population. Personal well-being in the latter case is $U(K/N_0)$. If the existing population were to have the misfortune of being too many ($N_0 > N^S$), the personal well-being of each of the N_0 individuals would be negative. And *that* is over-population.[42]

7.2 Coherence

A morality that requires people to award the same ethical weight to potential and present people, which is what Total Utilitarianism does, has been questioned by others. Narveson (1967), for example, had noted difficulties in the requirement and had recast Utilitarianism by observing that "we (utilitarians) are for making people happy, not for making happy people" (Narveson, 1973: 73). He called his theory Person-Affecting Utilitarianism.[43]

Parfit (1976) observed that Narveson's theory yields an incoherent notion of goodness. As his critique would apply equally to Generation-Relative Ethics, here is an example of what he, and in endorsing the criticism Broome (2004), meant.

We consider individual A and a potential person B, and apply the *ceteris paribus* clause on all other people. The prevailing state of affair (or "social state", to use the economist's term), *X*, is one where A's well-being is expected to be 11. But A has the option of creating B. Two social states that include B are possible. In one (we label it *Y*) A's and B's well-beings would each be 6. In the other (we label it *Z*) they would be 8 and 4, respectively. *Y* and *Z* differ by way of the extent to which A and B share resources, a feature of those states of affair that are taken into account by A when contemplating whether to create B. We re-label *X*, *Y*, and *Z* as {11}, {6,6}, and {8,4}. The problem is to rank them in a situation where A exists but B doesn't. We say that

---

[42] We are assuming that killing a fraction of the existing population so as to raise the well-being of those that remain to positive levels is forbidden.

[43] Narveson argued that we are under no obligation to procreate, even if, other things equal, it was guaranteed that the person born would have an outstandingly good life. Obligation is a strong word. In this paper I am exploring a particular way of evaluating states of affair, in which "well-being" serves as the coin with which states of affair are evaluated. Generation-Relative Utilitarianism is a lot closer to Total Utilitarianism than Narveson's theory. I have friends however who regard Generation-Relative Utilitarianism to be infested with ontological elitism. The criticism comes close to regarding potential persons as actual persons, a view that the problem of Sleeping Beauty asks us to reject.

B, were he to be created, would be of a different generation from A - hence "Generation-Relative Ethics".

Denote ethical dominance by ">" and ethical equivalence by "=". We imagine A to be a Generation-Relative Utilitarian. Suppose the weight A awards to the well-being of a potential person is half the weight awarded to the well-being of an existing person, the latter being set equal to 1. That means $\pi = 1/2$. It also means that so long as B is a potential person, $X > Z > Y$. But if B were born, he would be an actual person. The ethical relation between Y and Z, the only remaining social states, would then be $Y = Z$. That's Parfit's criticism.

The trouble with the criticism is that it harks back to the Genesis Problem. As the name suggests, the Genesis Problem is viewed by someone residing nowhere. That is why Total Utilitarianism would say $Y = Z > X$ before B is conceived. Parfit took exception to Narveson's Person-Affecting Utilitarianism because the binary relation between social states is dependent on whether B is an actual person. Generation-Relative Utilitarianism has that same feature. In the present example it is A who is evaluating social states. She does that from somewhere, she is not nowhere.

It is key to Generation-Relative Ethics that the ranking of the three states of affair X, Y, and Z is a function of the state of affair at which the ranking is arrived at. The perspective is from the current state. For example, the ethical content of the pair {6,6} in a world where B is only a potential person differs from the ethical content of the pair {6,6} in a world where B is an actual person. Denote the ranking of the triplet $\{X,Y,Z\}$ at X (the social state at which only A is present) under Generation-Relative Utilitarianism by the relationships "$>_X$" and "$=_X$", and define ("$>_Y$", "$=_Y$") and ("$>_Z$", "$=_Z$"), correspondingly, at Y and Z.

Because A subscribes to Generation-Relative Utilitarianism, she would choose not to add to the population, because viewed from X,

$$\{11\} >_X \{8,4\} >_X \{6,6\}. \tag{8}$$

If A were to view the triplet from Y, however, the ranking would be

$$\{6,6\} =_Y \{8,4\} >_Y \{11\}, \tag{9}$$

and from Z, it would be

$$\{6,6\} =_Z \{8,4\} >_Z \{11\}. \tag{10}$$

There is no contradiction here. Once X ceases to be an element of the set of possible outcomes (because B has come into existence), the ranking over Y and Z is reversed.[44]

[44] Total Utilitarianism and General-Relative Utilitarianism are not the same normative theory. The binary relations over social states that are implied by the two are different. We should therefore distinguish them. Under Total Utilitarianism the social state in which the ethics is conducted does not matter. The ranking of $\{X,Y,Z\}$ is the same ($Y = Z > X$) regardless of whether B exists. If a temporal structure were imposed on the example, X would not be reachable if either Y or Z were realized; nevertheless the thought experiment that would give rise to inequalities (8)-(10) could still be conducted. I am grateful to Simon Beard and Shamik Dasgupta for clarifying the complexities of the example.

8 Population Ethics Across the Generations

The "preference reversal" displayed in inequalities (8)-(10) translates to "intergenerational inconsistency", in that each generation's ordering over possible future worlds differs from those of future generations over the possible future worlds they in turn face. But as we confirm below, it is possible for the generations to arrive at choices that are consistent with one another's perspective. We now confirm how each generation would be able both to influence and affirm the choices to be made by those that are to follow. We do that by building on the timeless model we have just studied to construct a world moving through time.

8.1 The Model

Time is divided into periods and is denoted by t ($t = 0,1,...$). The horizon is indefinite, but there is a positive risk of extinction of the human race at t should it have survived until then. That rate is taken to be a constant, δ (> 0). δ operates like a constant time discount rate. Define $\theta = 1/(1+\delta)$. θ assumes the role of a "time discount factor". I imagine that other than humanity's survival there is no future uncertainty.[45]

It is imperative to distinguish stocks from flows in economies moving through time. Flows are the rates at which stocks change with time. For starkness we imagine the global economy and regard K to denote Nature. K is a stock, measured, say in units of biomass. To retain the parsimony of the timeless economy, we are to imagine that people are pure consumers, they themselves do not engage in production activities (but see the Appendix) and have no possibility of investing or dis-investing. Nature is assumed to do all the producing, providing humanity with a flow of a non-storable, all-purpose consumption good, of amount $F(K)$.

We assume that Nature isn't mishandled, meaning that K does not decline. I realize the latter assumption is a stretch, but I am trying to uncover the optimum population and consumption policy for the economy. It would make no sense to imagine that the economy's institutions are in a bad shape. Even otherwise the model is sensible. For example, in not allowing for the possibility of capital accumulation, the model encourages us to focus on the limits set by Nature: K captures environmental constraints. Mismanaging Nature would amount to reducing K. I elaborate on $F(K)$ in Section 9-10 when putting the theory to work on data on Nature's capacity to offer humanity with a flow of ecological services.

People live for two periods. In their first period (childhood) they are maintained by their parents. At the beginning of their second period (adulthood) - should humanity survive until then - they choose how many children to have and how to share $F(K)$ among themselves and their children. The way I have phrased matters could suggest that the society we are studying is

---

[45] δ describes a Poisson process. We may think of it as reflecting the possibility of a global pandemic, uninfluenced by policy. Rees (2003) offers an assessment of the likelihood of such a global disaster.

authoritarian and centralized. It isn't. Rather, we suppose people of the same generation are identical and that adults in any period have the same ethical motivation. So, even though it is individuals who do the choosing, we may speak as though each generation chooses the size of the next generation. Agent-based restrictions on choice translate here into generation-based restrictions.

People die at the end of their second period. There are thus two generations of people in any period. This is not good demography, but it turns out not to matter. To include realistic demographic features would not add to the substance of population ethics.

Let $N(t)$ be the number of adults at t, should the world survive until then. I will call them "generation-t". The number of children they choose to have is $N(t+1)$. The two generations at t make a total population of $N(t)+N(t+1)$, sharing $F(K)$. In period $t+1$, assuming the world survives, the $N(t+1)$ children of the previous generation are adults (they constitute generation-$(t+1)$); and if they give birth to $N(t+2)$ children, they share $F(K)$ with them. And so on.

Personal well-being is a function of her consumption level. Adults and children have the same well-being function, $U(C)$, and U has the same technical properties as in the timeless model we studied previously. It follows that generation-t has two choices to make: the size of the next generation ($N(t+1)$) and the basis on which to share $F(K)$ with them. In view of the properties of the U-function we have assumed, the latter choice is simple enough: parents share their resources equally with their children. Let $C(t)$ denote an individual's rate of consumption at t. Then $C(t) = F(K)/(N(t)+N(t+1))$. That's backward induction, a pre-requisite in Generation-Relative Utilitarianism.

The normative criterion each generation adopts for choosing the size of their family is "intergenerational well-being". How would Total Utilitarianism give expression to the notion? Here is Sidgwick (1907: 414) on the matter:

"It seems ... clear that the time at which a man exists cannot affect the value of his happiness from the universal point of view; and that the interests of posterity must concern a Utilitarian as much as those of his contemporaries, except in so far as the effect of his actions on posterity - *and even the existence of human beings to be affected* - must necessarily be more uncertain." (Italics added)

Let $V(t)$ denote intergenerational well-being as judged by generation-t. Total Utilitarianism regards $V(t)$ to be the (expected) sum of each person's well-being, from t into the indefinite future; and that is

$$\sum_{s=t}^{\infty}[N(s)U(F(K)/(N(s)+N(s+1)))\theta^{(s-t)}].$$[46]

[46] The discount factor θ reflects the risk of extinction in each period. There are other reasons one could offer for intergenerational discounting. The classic on this is Koopmans (1960, 1972). Dasgupta (2004, 2012) and Kelleher (2016) construct the philosophical underpinnings of those other reasons. I am taking a minimalist view here, by regarding the risk of extinction as the sole reason for discounting future well-beings.

Generation-Relative Utilitarianism in contrast judges $V(t)$ to be a weighted value of the expected sum of personal well-beings. The theory awards a weight π (< 1) to potential well-beings regardless of the date of arrival of potential people.

Because the economy has a stationary structure, generation-t can reason that future generations will have a motivation very much like theirs. It follows that

$$V(t) = N(t)U(F(K)/(N(t)+N(t+1))) + \pi N(t+1)U(F(K)/(N(t)+N(t+1))) + \pi\{_{s=t+1}\Sigma^{\infty}[\theta^{(s-t)}[N(s)U(F(K)/N(s)+N(s+1)) + N(s+1)U(F(K)/N(s)+N(s+1))]\},$$

$$0 < \pi, \theta < 1 \text{ and } t = 0,1,.... \qquad (11)$$

The present is taken to be $t = 0$.[47]

$V(t)$ guides the choices of generation-t, for all $t \geq 0$. Generation-0 is $N(0)$ in number. So $N(0)$ is given. If previous generations hadn't been moved by Generation-Relative Utilitarianism, the population size at $t = 0$ would not have been chosen with that theory as their guide. No doubt generation-0 will want to ask what their ancestors' intentions had been, and feel obliged to choose so as not to be entirely at variance with them; but those considerations would be a distraction here. So I assume that generation-0 regards the past to be past. They are entirely forward looking. Equation (11) binds members of generation-t to their descendants. As all generations from $t = 0$ onward adhere to Generation-Relative Utilitarianism, equation (11) is taken to bind members of all generations to their descendants.

8.2 Intergenerational Consistency

The analysis that follows is informal. Beyond supposing that $N(0)$ is small relative to K (otherwise generation-0 will choose not to reproduce), I avoid technicalities.

Consider an arbitrary t ($t \geq 0$). $N(t)$ will have been chosen by the previous generation. That means $N(t)$ will not be subject to choice at t. But $N(t+1)$ will be subject to choice by generation-t, comprising $N(t)$ people. It could seem that the optimum population numbers - from t onward - are those that together maximize $V(t)$. But there are two problems with that thought; one practical, the other ethical. The practical problem has to do with implementation. Generation-t has no means of enforcing its will on future generations. Moreover, later generations will have every incentive not to procreate in accordance with generation-t's bidding. Because $\pi < 1$, their grounds of binding reason will differ from that of generation-t, and in turn from those of one another. That's intergenerational inconsistency.

Turning to the ethical matter, it can be argued that generation-t would have no right to enforce its will on future generations even if a method was found to enforce it. Consider then a future generation, say, s ($s > t$). Intergenerational well-being as viewed by generation-s will be $V(s)$, and $V(s)$ is the ground of binding reason for generation-s. An ideal population profile under

---

[47] "$_{s=t+1}\Sigma^{\infty}$" denotes summation from $t+1$ to infinity. Notice that equation (11) reduces to equation (6) in the timeless model if there is no period following t. In that case only the first two terms in the right hand side of equation (11) remain. But that's equation (6).

Generation-Relative Utilitarianism is therefore a sequence of reproductive choices $\{N^*(1),...,N^*(t), N^*(t+1),...\}$ with the property that $N^*(t+1)$ maximizes $V(t)$ on the supposition that all generations following t will choose in accordance with that sequence. That common supposition has the status of an implicit understanding among the generations. Because each generation has an understanding of the choices its descendants will face (assuming there will be descendants!), it can anticipate the choices they will make. Its own choice is based on a forecast of its descendants' choices. That said, the exercise isn't exclusively forward looking. Each generation looks back at the choices their ancestors had made, if only to reassure itself of the implicit understanding and confirm that its own choice is guided by it. The population sequence $\{N^*(1),...,N^*(t+1),...\}$ is intergenerationally consistent, even though the ethical motivations of the generations are not consistent with one another.[48]

8.3 Extreme Theories

Total Utilitarianism represents one extreme set of values. At the other extreme is a theory that is easy to describe but is also questionable. It involves the thought that *no* weight should be awarded to potential well-beings. The viewpoint was explored by Enke (1966) in his study of social cost-benefit analysis of family planning programmes in poor countries. Enke sought ways to measure the economic value of prevented births, which he took to be the discounted sum of the differences between an additional person's consumption and the person's output over his lifetime. Children in Enke's theory have value only if they pay their way. In our formal model Enke's theory would amount to setting π equal to zero.

But Enke's is an extreme point of view, as is Total Utilitarianism: one awards no weight to potential well-being; the other awards the same weight to potential well-being as it does to the well-being of present or future people. We have reasons to reject both. Generation-Relative Utilitarianism lies between two extremes and reflects the strength of each without giving in to the weaknesses of either. It prescribes neither a large population nor a small population. Instead, it offers a wide space in between, within which more detailed ethical considerations can be embedded. We should not expect an ethical theory to do more.

8.4 The Stationary Optimum

Characterizing the implicit understanding among the generations for arbitrary values of $N(0)$ is hard work. So we study the understanding that supports a stationary profile. Along a stationary profile (we confirm presently that it is unique) the generations replicate their numbers. Imagine that N^* is that common number.

The idea underlying $\{N^*,N^*,...\}$ can be expressed in the following way: If $N(0)$ happens to equal N^*, replication would be the optimum choice of every generation if they were to suppose that all subsequent generations will in turn replicate their numbers. N^* is then a self-enforcing

[48] $\{N^*(1),...,N^*(t+1),...\}$ is an intergenerational Nash equilibrium.

reproductive policy for every generation. The understanding that implements the policy in question is only implicit of course. But because N^* is a self-enforcing choice for each generation, there is no call for a constitutional directive. We now identify the equation the population sequence $\{N^*,N^*,\dots\}$ must satisfy.[49]

In each period there are N^* adults and N^* children, making a total population of $2N^*$. In each period each person consumes $F(K)/2N^*$ units of the consumption good. Write $C^* = F(K)/2N^*$. It is now simple to show that C^* (correspondingly N^*) is the solution of the equation,[50]

$$dU(C)/dC = [2\pi(1+\theta)/(1+\pi+2\pi\theta)]U(C)/C. \quad (12)$$

Define

$$P = 2\pi(1+\theta)/(1+\pi+2\pi\theta). \quad (13)$$

Equation (12) can then be re-written in the compact form

$$U(C)/dC = PU(C)/C. \quad (14)$$

P represents the proportionate gap between marginal well-being and the average well-being per unit of consumption. The smaller is P, the greater is that gap. It is easy to confirm that $P \leq 1$.

Equation (14) is fundamental to our analysis here and extends Generation-Relative Utilitarianism beyond the timeless world (equation (7)) to cover one that moves through time.[51]

---

[49] If U satisfies equation (2), the stationary N^*-profile is locally stable. That's the ultimate justification for studying the stationary profile.

[50] To confirm that, we open up equation (11) and write it as

$$V(t) = N(t)U(K/(N(t)+N(t+1))) + \pi N(t+1)U(K/(N(t)+N(t+1))) + \pi\theta U(K/(N(t+1)+N(t+2))) + \pi\theta U(K/(N(t+1)+N(t+2)) + \dots, \quad 0 < \pi, \theta < 1.$$

As previously we suppose that $N(t+1)$ is a continuous variable. Differentiating $V(t)$ with respect to $N(t+1)$, setting the derivative equal to zero, and assuming that $N(t) = N(t+1) = N$ at all t yields equation (12).

[51] Equation (14) is the stationary equilibrium of a "game" involving the generations that are born. Each generation chooses the size of the generation that is to follow. Equation (11) summarizes the motivation of generation t, and by recursion the motivation of all generations that follow t. In more realistic models of individual life cycles parents' ethical motivations will not remain the same as and when they bear children. For example, adults in the world we know live for more than one period and choose the number of children they have sequentially. A simple exercise with Generation-Relative Utilitarianism confirms that a parent's ethical motivation before the birth of the first child differs from the motivation she would have after the birth of the first child. In Section 7.2 it was noted that the theory is constructed out of the view that the ethical ordering of states of affair isn't constructed from nowhere, but can depend on the state of affairs from which the ordering is constructed.

Shifting ethical perspectives were explored by Phelps and Pollak (1968), who studied savings behaviour in a world that is committed to Generation-Relative Utilitarianism but where future population sizes are not a matter choice. One consequence they drew is that the stationary equilibrium of their "savings game", where each generation chooses a saving rate, is inter-generationally inefficient; that is, there exists a feasible saving policy, one for each generation, that would result in an outcome where the level of well-being of every generation would be higher. (The superior saving policy would not be an equilibrium, however.) It would be

9 Earth's Carrying Capacity

Our intertemporal model has no production, no technological change, more broadly none of the pathways humanity has used to better ourselves over the millennia. That demands justification.

In Dasgupta (1969) Total Utilitarianism was studied in a model economy in which labour, accumulated capital, and a fixed quantity of land (for which read, "biosphere") produces output. But even if discoveries and new ideas were added to the list of factors, substitution possibilities between the produced factors and the biosphere would be increasingly limited if the latter were to dwindle in extent and quality.[52] The model in Dasgupta (1969) assumed considerable substitution possibilities (the elasticity of substitution was assumed to be 1). Nevertheless the finiteness of land was found to be a biting constraint. The average living standard and population size under Total Utilitarianism were shown to tend in the long run to finite levels. The simple consumption model I am using here mimics that feature.

9.1 The Biosphere as a Renewable Natural Resource

We are to imagine biosphere as a renewable natural resource (K), producing a flow of ecological services ($F(K)$), measured say in units of biomass. That amounts to a huge aggregation exercise, for Nature is a mosaic of degradable assets. Mismanagement of it amounts to reductions in K even when K is small relative to human needs. Agricultural land, forests, watersheds, fisheries, fresh water sources, estuaries, wetlands, the atmosphere - more generally, ecosystems - are assets that are self-regenerative, but can suffer from deterioration or depletion because of human encroachment. In a fundamental work on the state of the world's ecosystems, MEA (2005a-d) classified them in terms of the services they provide. The publications offered a four-way classification of ecosystem services: (i) provisioning services (food, fibre, fuel, fresh water); (ii) regulating services (protection against natural hazards such as storms; the climate system); (iii) supporting services (nutrient cycling, soil production); and (iv) cultural services (recreation, cultural landscapes, aesthetic or spiritual experiences). Notice that cultural services and a variety of regulating services (such as disease regulation) contribute directly to human well-being, whereas others (soil production) contribute indirectly (by providing the means of growing food

---

reasonable to imagine though that once the generations in the Phelps-Pollak world come to realize that the saving policy they are pursuing is inefficient, they would want to enrich the space of strategies from which they choose. Instead of choosing a saving rate (which is a number), each generation could in principle choose from saving policies (saving rates as functions of inherited wealth). Saving rates in equilibrium would then be the rates that are implied by equilibrium saving policies. Such a move can create equilibria that are efficient (Dasgupta, 1974b). I have not investigated how that move would work in the population problem.

[52] Ehrlich and Goulder (2007) has a fine discussion of this.

crops).[53]

To imagine the biosphere as a renewable natural resource requires facing a further problem. Even two thousand years ago, when global population was under 250 million and per capita income a bit over a dollar a day, it would have been a reasonable approximation to treat humanity as a separate entity from the biosphere. Today it is no longer possible to do so. We are much engaged in transforming the biosphere, by both creating biomass and destroying it. So we have to imagine humanity as at once a constituent of the biosphere and an entity separate from it. No doubt that's a stretch, but it is possible to do it without running into contradictions. We avoid contradiction by noting that a portion of $F(K)$, say α, is needed for the maintenance of the biosphere. So, if over a period of time $F(K)$ was to be usurped entirely by humanity, K would shrink and biodiversity would be reduced. If during an interval of time humanity was to consume even more than $F(K)$, K would shrink even more, further drawing down biodiversity. Humanity is doing that now, which is what has led Wilson (2016) to propose that we should leave half the biosphere (designed as a connected region) alone; that is, set α equal to 0.5. $(1-\alpha)F(K)$ should therefore be interpreted as the useable flow of biomass; useable, that is, by humanity.

It could be thought that $F(K)$ must be an increasing function of K for all values of K; but that would be to overlook that Earth is finite in extent. F should therefore be taken to be an increasing function of K for small values of K, but a declining function of K for large values of K. Earth's "carrying capacity" for the prevailing life-forms - a formidable notion in itself, but one that cannot be avoided - is that positive value of K at which $F(K)$ is zero.[54]

---

[53] Ecosystems combine the a-biotic environment with communities of plants, animals, fungi, and microorganisms to form functional units. The individual actors in an ecosystem include organisms that pollinate, decompose, filter, transport, re-distribute, scavenge, fix gases, and so on. Their activities enable ecosystems to maintain a genetic library, preserve and regenerate soil, fix nitrogen and carbon, recycle nutrients, control floods, mitigate droughts, filter pollutants, assimilate waste, pollinate crops, operate the hydrological cycle, and maintain the gaseous composition of the atmosphere. Among the visible products of ecosystems are food, fibres, fuel, and fresh water; but many of the products remain hidden from view.

[54] A tractable form of $F(K)$, in wide use among ecologists for a wide variety of ecosystems, is quadratic:

$F(K) = rK(1 - K/K^*)$, where r and K^* are positive constants.

In this equation r is the "intrinsic growth rate" of K (because r at small values of K is the percentage rate of growth of K) and K^* is the system's carrying capacity (because $F(K^*) = 0$).

The view that the biosphere is a renewable natural resource covers pollution as well (e.g. contemporary carbon emissions into the atmosphere). Pollutants are the reverse of natural resources. One way to conceptualize pollution is to view it as the depreciation of capital assets. Acid rains damage forests; carbon emissions into the atmosphere trap heat; industrial seepage and discharge reduce water quality in streams and underground reservoirs; sulfur emissions corrode structures and harm human health; and so on. The damage inflicted on each type of asset (buildings, forests, the atmosphere, fisheries, human health) should be interpreted as depreciation. For natural resources depreciation amounts to the difference between the aggregate rate at which it is harvested and its natural regenerative rate; for pollutants the depreciation they

Humanity's success in raising the standard of living over the past 250 years has in great measure involved mining and degrading *K*. That has usually gone unrecorded in official economic statistics, but we know now that even while industrial output increased by a multiple of 40 during the 20th century, the use of energy increased by a multiple of 16, methane-producing cattle population grew in pace with human population, fish catch increased by a multiple of 35 and carbon and sulphur dioxide emissions rose by more than 10. The application of nitrogen to the terrestrial environment from the use of fertilizers, fossil fuels, and leguminous crops is now at least as great as that from all natural sources combined. Soil nitrogen and phosphorus inventories have doubled over the past century (nitrate levels in Greenland ice are today higher than at any time in the previous 100,000 years), and 40-45 per cent of the 45-60 billion metric tons of carbon harnessed annually by terrestrial photosynthesis is now appropriated for human use (Vitousek et al., 1986, 1997). Extinction rates of species since the early modern era have been far above background rates and have increased a lot further since the 19th century (RSPB et al., 2013).

Although the rise in the concentration of atmospheric carbon receives much the greater public attention today, the Millennium Ecosystem Assessment (MEA, 2005a-d) reported that 15 of the 24 ecosystems the authors had investigated world-wide are either degraded or are being exploited at unsustainable rates. Population pressure on the land in use and habitat destruction that accompanies human encroachment are the proximate causes. The figures put the scale of humanity's presence on the planet in perspective and record that we are now Earth's dominant species (Ehrlich and Ehrlich, 2008). The statistics also explain why our epoch has now been renamed the Anthropocene.[55] These all point to rates of biomass transformation in excess of the useable flux, $(1-\alpha)F$. Consequently, they point to reductions in *K*.

9.2 Ecological Footprint

Studying biogeochemical signatures of the past 11,000 years, Waters, Zalasiewicz, Summerhayes et al. (2016) have provided a revealing sketch of the Anthropocene. The authors noted that a sharp increase took place in the middle of the 20th century in the rate of deterioration in the workings of Earth's life support system. They proposed that mid-20th Century should be

---

inflict on natural resources is the difference between the rate at which pollutants are discharged into the resource-base and the rate at which the resource-base is able to neutralize it. The task in either case is to estimate those depreciations. It follows that there is no reason to distinguish the analytical structure of resource management problems from pollution management problems. Roughly speaking, "resources" are "goods", while "pollutants" (the degrader of resources) are "bads". Pollution is the reverse of conservation.

[55] The term "Anthropocene" was popularized by Crutzen and Stoermer (2000) to mark a new epoch that began with the Industrial Revolution some 250 years ago. Wilson (1992, 2016) are two authoritative accounts of the value of bio-diversity and its precipitous decline over the past century.

regarded as the time we entered the Anthropocene.[56]

Their reading is consistent with macroeconomic statistics. World population in 1950 was 2.5 billion. Global GDP was a bit over 5 trillion dollars (PPP). The average person in the world was poor, with an annual income of a bit over 2,000 dollars (PPP). Since then the world has prospered materially beyond recognition. World income per capita today is 15,000 dollars (PPP) and population has increased to 7.3 billion. World output of final goods and services today is about 108 trillion dollars (PPP), which helps to explain not only the stresses to the Earth system that we have just reviewed, but also hints at the possibility (one that we confirm below using crude data) that humanity's extraction of biomass has for some time exceeded sustainable levels ($(1-\alpha)F$).

In a review of the state of the Earth's life support system, WWF (2008) reported that in the early years of this century, humanity's demand for ecological services exceeded by 50 per cent the rate at which the biosphere is able to supply those services to us. The figure is based on the idea of "global ecological footprint", which is the surface area of biologically productive land and sea needed to supply the resources a human population consumes (food, fibres, wood, water) and to assimilate the waste it produces (materials, gases). The Global Footprint Network (GFN) regularly updates their estimates of the global ecological footprint.[57] A footprint in excess of 1 means demand for ecological services exceeds their supply. That means to maintain the world's average living standard (roughly 12,000 international dollars in the early 2000s), we would have needed 1.5 Earths. This is, to be sure, a very, very rough figure; but they are the only estimates we have to date to work with. Ecology continues to be a vastly underfunded discipline.

GFN's most recent estimate is a footprint of a bit over 1.6, which in our terminology means humanity has in recent years been consuming ecological services at the rate $1.6(1-\alpha)F(K)$. Humanity's demand for ecological services can exceed supply for a period, but not indefinitely. Our model would interpret a footprint in excess of 1 as a decline in K (i.e. $\Delta K < 0$). Sustainable development would require that the footprint over time must on average equal 1. To be sure, $F(K)$ can be made to increase by measures that reduce the footprint to less than 1. For example, advances in bio-technology are designed to increase $F(K)$, but the advances would be successful only if they don't have large, unintended adverse consequences on the biosphere. Moreover, irreversible losses, arising say from biological extinctions, would act as constraints on the biosphere's ability to recover. That said, moves toward consumption and production practices that

[56] The Anthropocene Working Group has recently proposed that the immediate post-war years should be regarded as the start of the Anthropocene. See Vosen (2016).

[57] For pioneering work on the idea of ecological footprints, see Rees and Wackernagel (1994) and Rees (2001, 2006). See also Kitzes et al. (2008). Wakernagel, who founded the Global Footprint Network (www.footprintnetwork.org/public), was a lead author of WWF (2008).

make smaller demands on the biosphere would be a more direct approach to reducing our impact on the Earth system.

Our model errs in regarding $F(K)$ to be constant over time. Because my aim here is to study Generation-Relative Utilitarianism in a world with no institutional imperfections, we are to imagine that human institutions are strong enough to prevent the ecological footprint from exceeding 1. To determine what Generation-Relative Utilitarianism prescribes in a world where the footprint exceeds 1 we would need to reduce K, estimate $F(K)$ at the reduced value of K, and re-compute N^*, and thereby C^*. That's the route I take below and, to reduce notation, assume that $\alpha = 0$.

10 Quantitative Estimates of Optimum Population

Equation (13) says that $P = 1$ if $\pi = 1$. But if $P = 1$, equation (14) reduces to the Sidgwick-Meade Rule (equation (1)). That is so no matter what θ happens to be. The reason for the latter is simple: Because there is no production in the economy, the optimum policy in each period is independent of the risk of extinction.[58]

Equation (13) also says that $P < 1$ if $\pi < 1$. But $P < 1$ implies $dU(C)/dC < U(C)/C$. Personal well-being is greater under Generation-Relative Utilitarianism than under Classical Utilitarianism. Correspondingly, population is smaller.[59]

By how much? To investigate, we consider once again the U-function in equation (2). Well-being subsistence is $C^S = B^{-1/\sigma}$ and the system's carrying capacity is $2N^S = F(K)/C^S$. Equations (13)-(14) say that if C^* is the optimum stationary living standard, then

$$C^*/C^S = N^S/2N^* = [(\sigma + P)/P]^{1/\sigma} > 1. \qquad (15)$$

Equation (15) is the basis on which we can test Generation-Relative Utilitarianism.

It is interesting that C^* does not depend on $F(K)$; it depends only on the (ethical) parameters π and C^S. But N^S is a function of C^S and $F(K)$, and equation (15) says that N^S/N^* is independent of $F(K)$. That has important consequences for demographic and economic policy. Suppose it is found that $F(K)$ is lower that what it was previously thought to be. Generation-Relative Utilitarianism says that C^* should be fully protected and that the entire adjustment should be borne by the population size. The theory recommends that N^* is reduced in proportion to the decline in $F(K)$.

We now estimate C^* and N^* by putting equation (15) to work on contemporary global statistics.

10.1 Critical Parameters

---

[58] The case "$\theta = 1$" requires comment because $V(t)$ in that case is unbounded. In a classic paper Koopmans (1965) discovered a way to re-formulate the idea of optimum intergenerational well-being so as to overcome the problem (economists call the revised formulation the "overtaking criterion" of optimality). Koopmans' suggestion works here as well.

[59] θ is pertinent to the case $P < 1$; it influences N^*.

The stumbling block proves to be C^S and P.[60] Of the pair, P is possibly the less problematic. One can imagine P being estimated from surveys on fertility intentions at the level of the household. No doubt household behaviour is not the exclusive source of moral facts (Section 1 drew attention to a number of reasons why), but it would be impudent (not to say authoritarian) to ignore household intentions altogether in reaching ethical directives.

The sticking point will almost certainly be C^S. Matters aren't helped by the fact that C^S has been misinterpreted for over a century. Identifying C^S involves a value judgment, but it is also embedded in cultural presumptions. To make matters even more problematic, C^S is politically sensitive. The latter problem would be eased if we were able to relate C^S to politically acceptable parameters such as absolute poverty lines (e.g. the World Bank's "1.25 dollars-a-day", or the institution's bolder criterion, "2 dollars-a-day"). The analysis of Section 5 showed however that no such relationship exists. A correct reading of C^S says that it is a lot higher than what people today judge to be an absolute poverty line. An alternative would be to regard relative poverty as a surrogate for C^S. It could appear odd that well-being subsistence depends on a society's average level of income and is an increasing function of it; but as we noted previously, there is much evidence that personal well-being depends in part on our own standard of living when compared to the living standard of others in our peer group.[61]

The sad truth is that population ethics has been so comprehensively neglected in the social sciences that we have developed no informed intuition about either C^S or P. Economists and decision theorists have identified reasons for commending that σ should be in the range $(0,1]$, they have devised techniques for estimating income inequality, social discount rates, the cost of carbon in the atmosphere, and identifying both absolute and relative poverty lines; but on ethical parameters that are crucial to population ethics, we have developed little intuition.

10.2 Sensitivity Analysis

Sensitivity analysis is about the best we can do today. But sensitivity analysis alone is not helpful. It informs but doesn't tell us how to proceed from there. One thought is to iterate: (i) start with an arbitrary pair of figures for C^S and P and derive C^* and N^*; (ii) revise the figures and estimate the corresponding C^* and N^*; (iii) continue doing so until the evaluator reaches what Rawls would call a "reflective equilibrium" regarding C^* and N^*. Of course, a pair of figures for C^* and N^* that "feel" right may still be very wrong; but we would then want to know why they are wrong. That would force us to think on the matter and engage in discussions with others.

For simplicity of calculation, I suppose $\sigma = \theta = 1$. To illustrate sensitivity analysis, I consider three sets of figures for C^S and π (the latter yields P) and deduce the C^* and N^* that are

[60] I have set $\sigma = 1$ for computational ease.

[61] I am grateful to Robert Solow for drawing my attention to the significance of relative consumption for well-being subsistence.

implied by each pair. I start with "large" figures for C^S and π and reduce them in steps.

We first need an estimate for $F(K)$. Consider the world as a whole. Annual global income today is about 108 trillion international dollars.[62] Even though GFN's most recent estimate of the ecological footprint is 1.6, I use the conservative figure of 1.5. Making the simplifying assumption that GDP draws proportionately on ecosystem services at all levels, we conclude that the sustainable level of annual world income is 72 trillion (2x108/3 trillion) dollars. That gives us $F(K)$.

We now study three illustrative cases:

(1) Suppose $\pi = 0.1$. Then $P \approx 0.325$ and $(\sigma+P)/P \approx 4$. Now suppose $C^S = 7{,}200$ dollars, which is approximately Guatemala's per capita GDP. It follows that $N^S = 10$ billion (72 trillion/7.2 thousand). Now use equation (14) to confirm that $2N^* = 2.5$ billion and $C^* = 28{,}800$ dollars. The latter figure is the per capita GDP of Portugal. Total Utilitarianism ($\pi = 1$) in contrast would arrive at a figure that is half as much (14,400 dollars) and a population size that is twice as much (5 billion). For easy comparison, I note that world population today is 7.2 billion and average world income per person is 15,000 dollars.[63]

(2) Suppose $\pi = 0.05$. Then $P \approx 0.18$ and $(\sigma+P)/P \approx 6.5$. Now suppose $C^S = 6{,}000$ dollars, which is roughly the per capita GDP of the Republic of Congo today. Then $N^S = 12$ billion (72 trillion/6 thousand). We can now use equation (14) to confirm that $2N^* = 1.85$ billion and $C^* = 39{,}000$ dollars. The latter figure is the per capita GDP of France. In contrast Classical Utilitarianism would advocate a population of 6 billion, enjoying a living standard of 12,000 dollars.

(3) Suppose $\pi = 0.01$. Then $P \approx 0.04$ and $(\sigma+P)/P = 26$. Now suppose $C^S = 1{,}000$ dollars, which is just a bit over the per capita GDP of Niger - one of the world's poorest countries. Then $N^S = 72$ billion (72 trillion/1 thousand); and so $2N^* = 2.8$ billion and $C^* = 26{,}000$ dollars. The latter is a bit above per capita GDP in Greece. In contrast Classical Utilitarianism would advocate a population of 36 billion at a standard of living of 2,000 dollars.

Although three exercises can't reveal much, they reveal that Generation-Relative Utilitarianism is most distinct from Classical Utilitarianism when π is small (and thereby P is small) and simultaneously C^S is small. I expected the former but didn't expect it would come tied to the latter.

11 Existential Risks and Rational Ends

---

[62] International dollars are a fictitious currency that evens out differences in the cost of living across countries. International dollars are often expressed as dollars adjusted for purchasing power parity (PPP). In what follows I drop the qualifier and use the term "dollar".

[63] To confirm the latter, set $\theta = \sigma = \pi = 1$ in equation (14). It could appear odd that Classical Utilitarianism advocates both smaller population and lower living standard than is the case in the contemporary world, but recall that our calculations pertain to a sustainable Earth (K). At a world GDP of 108 trillion dollars humanity's living standard is not sustainable.

In a deep meditation on the significance of a possible nuclear holocaust in which humanity suffers extinction, the late Jonathan Schell distinguished two types of death:

"It is of the essence of the human condition that we are born, live for a while, and then die... But although the untimely death of everyone in the world would in itself be an unimaginably huge loss, it would bring with it a separate, distinct loss that would be in a sense even huger - the cancellation of all future generations of human beings." (Schell, 1982: 114-115)

Schell's book was originally published as a three-part essay in *The New Yorker* in 1981, at a time the Cold War had created an especial chill. Schell was a writer, not a professional philosopher, but he made not one false move in philosophical reasoning in the crucial middle chapter, Second Death. Both Total Utilitarianism and Generation-Relative Utilitarianism measure the loss from the Second Death in terms of the well-being of all who would not exist on account of human extinction. In the previous section we saw how those losses could be estimated.[64] Schell however made a radically different move. He spoke of the loss each of us alive today would suffer if we were to discover that there will be no one after we are gone. He located that loss not to any attachment we may have to humanity writ large, but to a devaluation of our own lives. And he used the artist and his art to make the point:

"There is no doubt that art, which breaks into the crusted and hardened patterns of thought and feeling in the present as though it were the prow of the future, is in radically altered circumstances if the future is placed in doubt. The ground on which the artist stands when he turns to his work has grown unsteady beneath his feet." (Schell, 1982: 163)[65]

Schell spoke of the artist, but he could have made the same case for all who create ideas and objects. Future people add value to the creators' lives by making their creations durable. Here the fact of a general assumption that people desire to have children is significant. An artist may regard his work to be far more important than parenting, but he is helped by the presumption that there will be future generations to bestow durability to his work.

The examples Schell pointed to were works of art and discoveries in the sciences. Those creations are public goods, and most people don't have the talent to produce them. Confining attention to public goods is not only limiting, it also raises an ethical dilemma: Suppose we all were indifferent to having children and stared only at the prospective costs of raising them. We would then free-ride, and the artist would be mistaken in his assumption that there will be future people to give durability to his work.

Nevertheless, the direction Schell was pointing to is exactly right. Public goods aren't the

---

[64] Sikora (1978) is an early exploration for a response to the question using Classical Utilitarianism as the guide.

[65] In a wide-ranging discussion of the ethical implications of global extinction Scheffler (2013) has also observed that our own lives would be diminished if there were to be no future people.

only objects of ethical significance. Our values and practices are significant too. Many are private, even confined to the family, and it is important to us that they are passed down the generations. Procreation is a means of making one's values and practices durable. We imbue our children with values we cherish and teach them the practices we believe are right not merely because we think it is good for them, but also because we desire to see our values and practices survive. Those values and practices are not public goods. On the contrary, we cherish them *because* they are intimate. They are stories we tell our children about their grandparents' foibles, of our own joys, sorrows, and discomfiture, and we instruct them on the family rituals we ourselves inherited from our parents. It seems to me our descendants do something supremely important for us: they add value to our lives that our own mortality would otherwise deprive them of. That is the reason we would not practise reproductive free-riding even if we found reproduction to be personally costly. The springs that motivate humankind to assume parenthood are deep and abiding. Their genetic basis explains the motivation, but doesn't justify it. Justification is found elsewhere. Our children provide us with a means of self-transcendence, the widest avenue open to us of living through time, not merely in time. Mortality threatens to render the achievements of our life transitory, and this threat is removed by procreation. The ability to leave descendants enables us to invest in projects that will not cease to have value once we are gone, projects that justify life rather than merely serve it. Alexander Herzen's remark, that human development is a kind of chronological unfairness because those who live later profit from the labour of their predecessors without paying the same price; and Kant's view that it is disconcerting that earlier generations should carry their burdens only for the sake of the later ones, that only the last should have the good fortune to dwell in the completed building - or in other words, that we can do something for posterity but it can do nothing for us - are a reflection of an extreme form of alienation.[66]

The motivation transmutes from the individual to the collective. Every generation is a trustee of the wide range of capital stocks (be it cultural or moral, manufactured or natural) it has inherited from the past. Looking backward, it acknowledges an implicit understanding with the previous generation, of receiving the capital in return for its transmission, modified suitably in the light of changing circumstances and increasing knowledge. Looking forward, it offers an implicit proposal to the next generation, of bequeathing its stocks of capital that they in turn may be modified suitably by it and then passed on to the following generation. As was noted in our account of population ethics in a world moving through time, this perspective is not at odds with either Total or Generation-Relative Utilitarianism. Generation-t is moved to internalize the potential well-being of their descendants, as expressed in $V(t)$. Our descendants are not us, but

[66] Rawls (1972: 291) has a characteristically profound criticism of Kant's perspective.

they are not outside us either.[67]

Schell's reflections point also to the intrinsic value of Nature. It's a mistake to seek justification for the preservation of ecological diversity, or more narrowly the protection of, say, the giant redwoods, solely on instrumental grounds; that is on grounds that we know they are useful to us or may prove useful to our descendants. Such arguments have a role but they are not sufficient. Nor can the argument rely on the welfare of the members of such species (it does not account for the special role that species preservation plays in the argument), or on the "rights" of animals. A full justification bases itself as well on how we see ourselves, on what our rational desires are. In examining our values and thus our lives, we should ask if the destruction of an entire species-habitat for some immediate gratification is something we can live with comfortably. The idea of intergenerational exchange is embedded in the perspective of eternity, but the intellectual source of such exchange is a far cry from the conception that balked Herzen in his effort to locate mutually beneficial terms of trade. The mistake is to see procreation and ecological preservation as matters of personal and political morality. They are at least as much a matter of personal and political ethics.

[67] The idea of stewardship would appear to be common among different cultures. On African conceptions of intergenerational ethics, Behrens (2012: 189) writes: "African thought does not limit moral considerability to only the current generation. It conceives of a web of life that transcends generations, and of the environment as a resource shared by different generations. This entails a direct moral obligation to preserve the environment for future persons, since it is a communal good. Africans also expect that the current generation should develop an attitude of gratitude towards their predecessors for having preserved the environment on their behalf. This virtue of gratefulness ought to be realized by the current generation seeking to reciprocate by preserving the environment for future generations, in turn." I am grateful to Simon Beard for this reference.

Appendix

Generation-Relative Utilitarianism with Production

We consider a world where people are both producers and consumers. To do that we extend the model studied in Section 8.1 and reinterpret Nature to be a capital good. Human labour in combination with Nature produces a single consumption good.

Only adults are assumed to be capable of offering labour services. An individual therefore consumes in both periods but produces only when an adult. We assume that if $N(t)$ is the number of adults at t, the output of the consumption good is

$$Y(t) = K[N(t)]^{\beta} \qquad 0 < \beta < 1. \tag{A1}$$

As all output is consumed, consumption per head is obtained from equation (A1) as

$$C(t) = K[N(t)]^{\beta}/[N(t)+N(t+1)]. \tag{A2}$$

Inserting this consumption structure in the model studied in Section 8.1, it is simple to confirm that the optimum stationary consumption (and the corresponding population) sequence satisfies

$$\mathrm{d}U(C)/\mathrm{d}C = P^{**}U(C)/C, \tag{A3}$$

where

$$P^{**} = 2\pi(1+\theta)/(1+\pi+2\pi\theta(1-2\beta)). \tag{A4}$$

Comparing equations (13) and (A4), we note that $0 < P < P^{**}$. That means other things equal the proportionate gap between marginal well-being and average well-being per unit of consumption is greater in a world where people are producers and consumers, not simply consumers. Assume again that $U(C)$ satisfies equation (2). Equation (A3) then reduces to

$$C/C^{S} = N^{S}/2N = [(\sigma + P^{**})/P^{**}]^{1/\sigma} > 1. \tag{A5}$$

Let C^{**}, and correspondingly N^{**}, be the solutions of equation (A5). To obtain a feel for the numbers that are involved, suppose as in the first of our illustrations in the sensitivity analysis of Section 10.2 that $\pi = 0.1$, $\theta = 1$ and $\sigma = 1$. There are any number of estimates in the literature on labour's contribution to output as measured by its elasticity in production. In equation (A1) the elasticity is β. P^{**} in equation (A4) is a function of β. For illustration we, not unreasonably, assume $\beta = 0.5$. It follows that $P^{**} \approx 0.364$ and $(\sigma+P^{**})/P^{**} \approx 4$. To a first approximation production doesn't make much difference if β is in the neighbourhood of 0.5. Suppose, as in the previous illustration, that $C^{S} = 7{,}200$ dollars. Optimum population ($2N^{**}$) is, as previously, 2.5 billion and optimum consumption per capita (C^{**}) is, again as previously, 28,800 dollars (PPP). Both are far from where humanity is today.

References

Adams, R.M. (1979), "Existence, Self-Interest, and the Problem of Evil," *Nous*, 13(1), 53-65.

Arrow, K.J. (1965), *Aspects of the Theory of Risk-Bearing* (Helsinki: Yrjö Jahnssonin säätiö).

Arrow, K.J. (1981), "Optimal and Voluntary Income Distribution," in S. Rosefielde, ed., *Economic Welfare and the Economics of Soviet Socialism: Essays in Honor of Abram Berson* (Cambridge: Cambridge University Press).

Arrow, K.J. (1999), "Discounting, Morality, and Gaming," in P.R. Portney and J.P. Weyant, eds., *Discounting and Intergenerational Equity* (Washington, DC: Resources for the Future).

Arrow, K.J. and P.S. Dasgupta (2009), "Conspicuous Consumption, Inconspicuous Leisure," *Economic Journal*, 119(541), F497-F516.

Atkinson, A.B. (1970), "On the Measurement of Inequality," *Journal of Economic Theory*, 2(3), 244-263.

Behrens, K.G. (2012), "Moral Obligations Towards Future Generations in African Thought," *Journal of Global Ethics*, 8(2-3), 179-191.

Blackorby, C. and D. Donaldson (1982), "Ratio-Scale and Translation-Scale Full Interpersonal Comparability Without Domain Restrictions: Admissible Social-Evaluation Functions," *International Economic Review*, 23(2), 249- 268.

Blackorby, C. and D. Donaldson (1984), "Social Criteria for Evaluating Population Change," *Journal of Public Economics*, 25(1-2), 13-34.

Blackorby, C., W. Bossert, and D. Donaldson (1997), "Critical-Level Utilitarianism and the Population-Ethics Dilemma," *Economics and Philosophy*, 13(2), 197-230.

Bledsoe, C. (1994), "Children are Like Bamboo Trees: Potentiality and Reproduction in Sub-Saharan Africa," in K. Lindahl-Kiessling and H. Landberg, eds., *Population, Economics Development and the Environment* (Oxford: Oxford University Press).

Bongaarts, J. and J. Casterline (2013), "Fertility Transition: Is Sub-Saharan Africa Different?," *Population an Development Review*, 39(1), 153-168.

Boonin, D. (2014), *The Non-Identity Problem and the Ethics of Future People* (New York: Oxford University Press).

Brander, J.A. and M.S. Taylor (1998), "The Simple Economics of Easter Island: A Ricardo-Malthus Model of Renewable Resource Use," *American Economic Review*, 88(1), 119-138.

Broome, J. (1996), "The Welfare Economics of Population," *Oxford Economic Papers*, 48(2), 177-93.

Broome, J. (1999), *Ethics Out of Economics* (Cambridge: Cambridge University Press).

Broome, J. (2004), *Weighing Lives* (Oxford: Oxford University Press).

Caldwell, J.C. (1981), "The Mechanisms of Demographic Change in Historical Perspective," *Population Studies*, 35(1), 5-27.

Caldwell, J.C. (1982), *The Theory of Fertility Decline* (New York: Academic Press).

Cochrane, S.H. and S.M. Farid (1989), "Fertility in Sub-Saharan Africa: Analysis and Explanation," World Bank Discussion Paper No. 43, World Bank, Washington DC.

Cowen, T. (1989), "Normative Population Theory," *Social Choice and Welfare*, 6(1), 33-44.

Cowie, C. (2015), "Background Psychology in Non-Identity Cases." Unpublished Notes, Fitzwilliam College, Cambridge.

Crutzen, P.J. and E.F. Stoermer (2000), "The Anthropocene," *Global Change Newsletter*, 41, 17-18.

Daly, M.C., A.J. Oswald, D. Wilson, and S. Wu (2011), "Dark Contrasts: The Paradox of High Rates of Suicide in Happy Places," *Journal of Economic Behaviour & Organization*, 80(3), 435-442.

Dasgupta, A. and P. Dasgupta (2016), "Socially Embedded Preferences, Environmental Externalities, and Reproductive Rights," Under preparation.

Dasgupta, P. (1969), "On the Concept of Optimum Population," *Review of Economic Studies*, 36(3), 295-318.

Dasgupta, P. (1974a), "On Optimum Population Size," in A. Mitra, ed., *Economic Theory and Planning: Essays in Honour of A.K. Dasgupta* (Delhi: Oxford University Press).

Dasgupta, P. (1974b), "On Some Alternative Criteria for Justice Between Generations," *Journal of Public Economics*, 3(4), 405-423.

Dasgupta, P. (1989), "Population Size and the Quality of Life," *Proceedings of the Aristotelian Society*, Supp. Vol. LXIII, 23-40.

Dasgupta, P. (1993), *An Inquiry into Well-Being and Destitution* (Oxford: Clarendon Press).

Dasgupta, P. (1994), "Savings and Fertility: Ethical Issues," *Philosophy & Public Affairs*, (Spring), 99-127.

Dasgupta, P. (2005), "Regarding Optimum Population," *Journal of Political Philosophy*, 13(4), 414-442. Reprinted in J. Fishkin and R. Goodin, eds., *Population and Political Theory - Philosophy, Politics and Society, 8* (Chichester: Wiley-Blackwell), 2010.

Dasgupta, P. (2005), "Three Conceptions of Intergenerational Justice," in H. Lillehammer and D.H. Mellor, eds., *Ramsey's Legacy* (Oxford: Clarendon Press), pp. 149-169.

Dasgupta, P. (2008), "Discounting Climate Change," *Journal of Risk and Uncertainty*, 37(2-3), 141-169.

Dasgupta, P. (2012), "The Ethics of Intertemporal Distribution in a Warming Planet: Reply and Response to John E. Roemer," *Environmental & Resource Economics*, 50(4), 475-493.

Dasgupta, P. (2010), "The Place of Nature in Economic Development," in D. Rodrik and M. Rosenzweig, eds, *Handbook of Development Economics, Vol. V* (Amsterdam: North

Holland).

Dasgupta, P. and P. Ehrlich (2013), "Pervasive Externalities at the Population, Consumption, and Environment Nexus," *Science*, 340 (April 19), 324-328.

Edgeworth, F.Y. (1881), *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences* (London: Kegan Paul).

Ehrlich, P.R. and A.H. Ehrlich (2008), *The Dominant Animal: Human Evolution and the Environment* (Washington DC: Island Press).

Ehrlich, P.R. and L.H. Goulder (2007), "Is Current Consumption Excessive? A General Framework and Some Indications for the United States," *Conservation Biology*, 21(5), 1145-1154.

Ehrlich, P.R. and J.P. Holdren (1971), "Impact of Population Growth," *Science*, 171 (3977), 1212-1217.

Enke, S. (1966), "The Economic Aspects of Slowing Population Growth," *Economic Journal*, 76(1), 44-56.

Fortes, M. (1978), "Parenthood, Marriage and Fertility in West Africa," *Journal of Development Studies*, 14 (Special Issue on Population and Development), 121-149.

Galor, O. (2011), *Unified Growth Theory* (Princeton, NJ: Princeton University Press).

Goody, J. (1976), *Production and Reproduction* (Cambridge: Cambridge University Press).

Gosseries, A. (2010), "On Future Generations' Future Rights," in J. Fishkin and R. Goodin, eds., *Population and Political Theory - Philosophy, Politics and Society, 8* (Chichester: Wiley-Blackwell).

Gottlieb, M. (1945), "The Theory of Optimum Population for a Closed Economy", *Journal of Political Economy*, 53(3), 289-316.

Griffin, J. (1986), *Well-Being: Its meaning, measurement, and moral importance* (Oxford: Oxford University Press).

Guyer, J.L. (1994), "Lineal Identities and Lateral Networks: the Logic of Polyandrous Motherhood," in C. Bledoe and G. Pison, eds., *Nupitality in Sub-Saharan frica: Contemporary Anthropological and Demographic Perspectives* (Oxford: Clarendon Press).

Harford, J.D. (1998), "The Ultimate Externality," *American Economic Review*, 88(1), 260-265.

Harsanyi, J. (1955), "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Welfare," *Journal of Political Economy*, 63(4), 309-321.

Haveman, R. and B. Wolf (1995), "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature*, 33(4), 1829-1878.

Helliwell, J.F., R. Layard, and J.D. Sachs, eds. (2013), *World Happiness Report 2013* (New York: Earth Institute).

Heyd, D. (1992), *Genethics: The Morality of Procreation* (Berkeley, CA: University of

Calfornia Press).

Huemer, M. (2008), "In Defense of Repugnance," *Mind*, 117(468), 899-933.

Hurka, T. (1983), "Value and Population Size," *Ethics*, 93(4), 497-507.

Jamison, D.T., L.H. Summers et al. (2013), "Global Health 2035: a world converging within a generation," *Lancet*, 382 (December 7), 1898-1955.

Kavka, G.S. (1982), "The Paradox of Future Generations," *Philosophy & Public Affairs*, 11(2), 93-112.

Keheller, J. Paul (2016), "Pure Discounting in Climate Change Economics," *Economics and Philosophy*, forthcoming.

Kitzes, J., M. Wakernagel, J. Loh, A Peller, S. Goldfinger, D. Cheung, and K. Tea (2008), "Shrink and Share: Humanity's Present and Future Ecological Footprint," *Philosophical Transactions of the Royal Society, B*, 363(25 July), 467-475.

Kolm, S.-Ch. (1969), "The Optimal Production of Social Justice," in J. Margolis and H. Guitton, eds., *Public Economics* (London: Macmillan).

Koopmans, T.C. (1960), "Stationary Ordinal Utility and Impatience", *Econometrica*, 28(2), 287-309.

Koopmans, T.C. (1965), "On the Concept of Optimal Economic Growth," *Pontificiae Academiae Scientiarum Scripta Varia, 28*. Reprinted in T.C. Koopmans, *The Econometric Approach to Development Planning* (Amsterdam: North Holland), 1966.

Koopmans, T.C. (1972), "Representation of Preference Orderings over Time", in C.B. McGuire and R. Radner, eds., *Decision and Organization* (Amsterdam: North Holland).

Maddison, A. (2001), *The World Economy: A Millennial Perspective* (Paris: OECD).

McMahan, J.A. (1981), "Problems of Population Theory," *Ethics*, 92(1), 96-127.

MEA - Millennium Ecosystem Assessment - eds., R. Hassan, R. Scholes, and N. Ash - (2005a), *Ecosystems and Human Well-Being, I: Current State and Trends* (Washington, DC: Island Press).

MEA - Millennium Ecosystem Assessment - eds., S.R. Carpenter, P.L. Pingali, E.M. Bennet, and M.B. Zurek (2005b), *Ecosystems and Human Well-Being, II: Scenarios* (Washington, DC: Island Press).

MEA - Millennium Ecosystem Assessment - eds., K. Chopra, R. Leemans, P. Kumar, and H. Simmons - (2005c), *Ecosystems and Human Well-Being, III: Policy Responses* (Washington, DC: Island Press).

MEA - Millennium Ecosystem Assessment - eds., D. Capistrano, C. Samper K., M.J. Lee, and C. Randsepp-Hearne - (2005d), *Ecosystems and Human Well-Being, IV: Multiscale Assessments* (Washington, DC: Island Press).

Meade, J.E. (1955), *Trade and Welfare* (Oxford: Oxford University Press).

Nagel, T. (1979), "Death," in *Mortal Questions* (Cambridge: Cambridge University Press). Originally published in *Nous*, IV(1), February 1970.

Narveson, J. (1967), "Utilitarianism and New Generations," *Mind*, 76(1), 62-72.

Narveson, J. (1973), "Moral Problems of Population," *Monist*, 57(1), 62-86.

Ng, Y-K. (1986), "Social Criteria for Evaluating Population Change: An Alternative to the Blackorby-Donaldson Criterion," *Journal of Public Economics*, 29(3), 375-381.

Ng, Y.-K. (1989), "What Should We Do About Future Generations? Impossibility of Parfit's Theory X," *Economics and Philosophy*, 5(4), 235-253.

Parfit, D. (1976), "On Doing the Best for Our Children," in M. Bayles, ed., *Ethics and Population* (Cambridge, MA: Schenkman). Reprinted in J.S. Fishkin and R.E. Goodin, eds., *Population and Political Theory*, Philosophy, Politics and Society 8 (Chichester: Wiley-Blackwell), 2010.

Parfit, D. (1982), "Future Generations: Further Problems," *Philosophy & Public Affairs*, 11(2), 113-172.

Parfit, D. (1984), *Reasons and Persons* (Oxford: Oxford University Press).

Parfit, D. (2016), "Can We Avoid the Repugnant Conclusion?" *Theoria*, 82(2).

Phelps, E.S. and R. Pollak (1968), "Second-Best National Savings and Game Equilibrium Growth," *Review of Economic Studies*, 35(2), 185-199.

Pratt, J.A. (1964), "Risk Aversion in the Small and in the Large," *Econometrica*, 32(1-2), 122-136.

Ramsey, F.P. (1928), "A Mathematical Theory of Saving," *Economic Journal*, 38(4), 543-559.

Rawls, J. (1972), *A Theory of Justice* (Oxford: Oxford University Press).

Rees, M.J. (2003), *Our Final Century* (London: Random House).

Rees, W.E. (2001), "Ecological Footprint, Concept of," in S.A. Levin, ed., *Encyclopedia of Biodiversity, Vol. 2* (New York: Academic Press).

Rees, W.E. (2006), "Ecological Footprints and Biocapacity: Essential Elements in Sustainability Assessment," in J. Dewulf and H.V. Langenhove, eds., *Renewable-Based Technology: Sustainability Assessment* (Chchester, UK: John Wiley & Sons).

Rees, W.A. and M. Wackernagel (1994), "Ecological Footprints and Appropriated Carrying Capacity: Measuring the Natural Capital Requirements of the Human Economy," in A.M. Jansson *et al*., *Investing in Natural Capital: The Ecological Economics Appropriate for Sustainability* (Washington, DC: Island Press).

RSPB et al. (2013), *State of Nature 2013* - Report issued jointly by the Royal Society for the Protection of Birds and 19 other UK conservation and research associations - (London: RSPB).

Scheffler, S. (1994), *The Rejection of Consequentialism: A Philosophical Investigation of the Consideration Underlying Rival Moral Conceptions* (Oxford: Clarendon Paperback).

Scheffler, S. (2013), *Death & the Afterlife* (Oxford: Oxford University Press).

Schell, J. (1982), *The Fate of the Earth* (New York: Avon).

Schwartz, T. (1978), "Obligations to Posterity," in R. Sikora and B. Barry, eds., *Obligations to Future Generations* (Philadelphia: Temple University Press).

Seabright, P. (1989), "Population Size and the Quality of Life, II: Creating Persons," *Proceedings of the Aristotalian Society*, Supp. Vol. LXIII, 41-54.

Sen, A. (1970), *Collective Choice and Social Welfare* (San Francisco: Holden Day).

Sidgwick, H. (1907), *The Methods of Ethics* (London: MacMillan), 7th Edition.

Sikora, R.I. (1978), "Is It Wrong to Prevent the Existence of Future Generations?" in R.I. Sikora and B. Barry, eds., *Obligations to Future Generations* (Philadelphia: Temple University Press).

Sikora, R.I. (1981), "Classical Utilitarianism and Parfit's Repugnant Conclusion: A Reply to McMahan," *Ethics*, 92(1), 128-133.

Sterba, J.P. (1987), "Explaining Asymmetry: A Problem for Parfit," *Philosophy and Public Affairs*, 16(2), 188-192.

Temkin, L.S. (1987), "Intransitivity and the Mere Addition Paradox," *Philosophy and Public Affairs*, 16(2), 138-187.

Veblen, T. (1925, [1899]). *The Theory of the Leisure Class: An Economic Study of Institutions*. Originally published in 1899. Reprinted (London: George Allen & Unwin), 1925.

Vitousek, P.M., P.R. Ehrlich, A.H. Ehrlich, and P. Matson (1986), "Human Appropriation of the Product of Photosynthensis," *BioScience*, 36, 368-373.

Vitousek, P.M., H.A. Mooney, J. Lubchenco, and J.M. Melillo (1997), "Human Domination of Earth's Ecosystem," *Science*, 277, 494-499.

Vosen, P. (2016), "Anthropocene Pinned Down to Post War Period," *Science*, 353(6302), 852-853.

Waters, C.N., J. Zalasiewicz, C. Summerhayes, *et al*. (2016), "The Anthropocene is Functionally and Stratigraphically Distinct from the Holecene," *Science*, 351(6269), aad2622-(1-10).

Williams, B. (1993), *Shame and Necessity* (Berkeley, CA: University of California Press).

Wilson, E.O. (1992), *The Diversity of Life* (Cambridge, MA: Harvard University Press).

Wilson, E.O. (2016), *Half-Earth: Our Planet's Fight for Life* (New York: Liveright Publishing Corporation).

World Bank (2012), *World Development Indicators* (Washington DC: World Bank).

WWF - World Wildlife Fund - (2008), *Living Planet Report 2008* (Gland: WWF International).

Table 1

Social Statistics from Rich and Poor Regions (Year 2010)

	rich	poor
Population (millions)	1,127	796
GDP per capita (international dollars)	37,000	1,300
Total fertility rate*	1.8	4.1
Under-5 mortality rate (per 1000)	6	108
Life expectancy at birth (years)	80	59
Percentage enrollment in secondary education	100	39

Source: World Bank (2012)

* Total fertility rate (TFR) is the number of successful pregnancies that a woman expects to have during her reproductive years. The number 2.1 is usually taken to be the TFR that, over the long run would lead to a stable population.

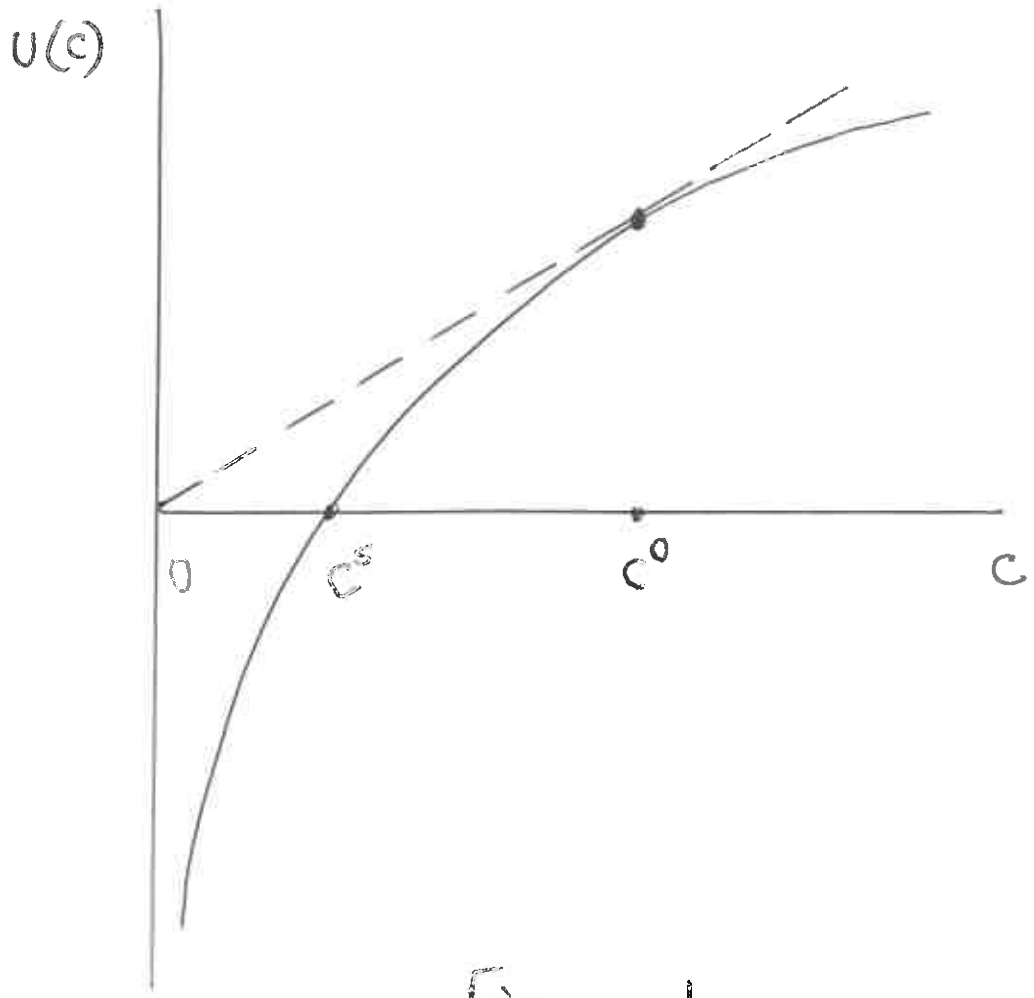


Figure 1

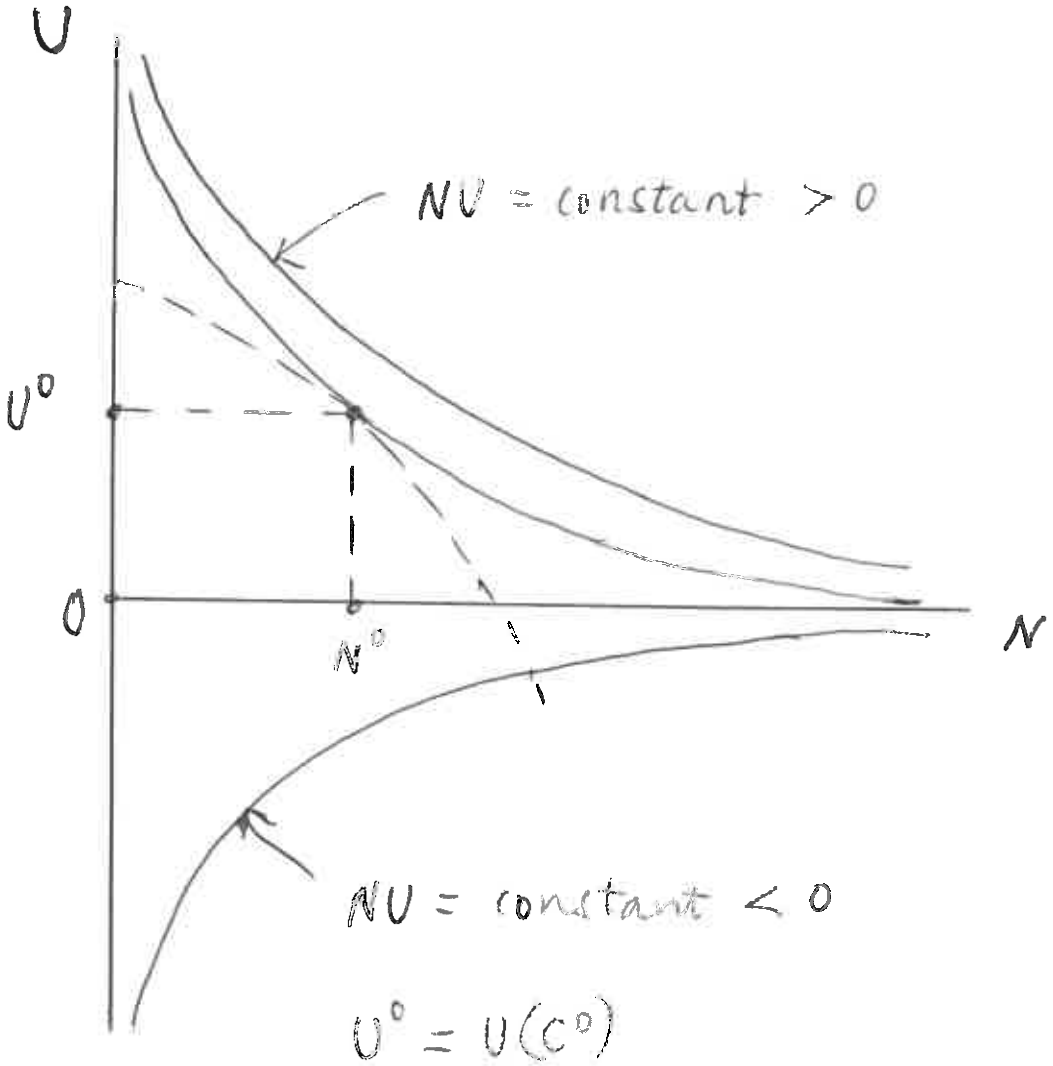
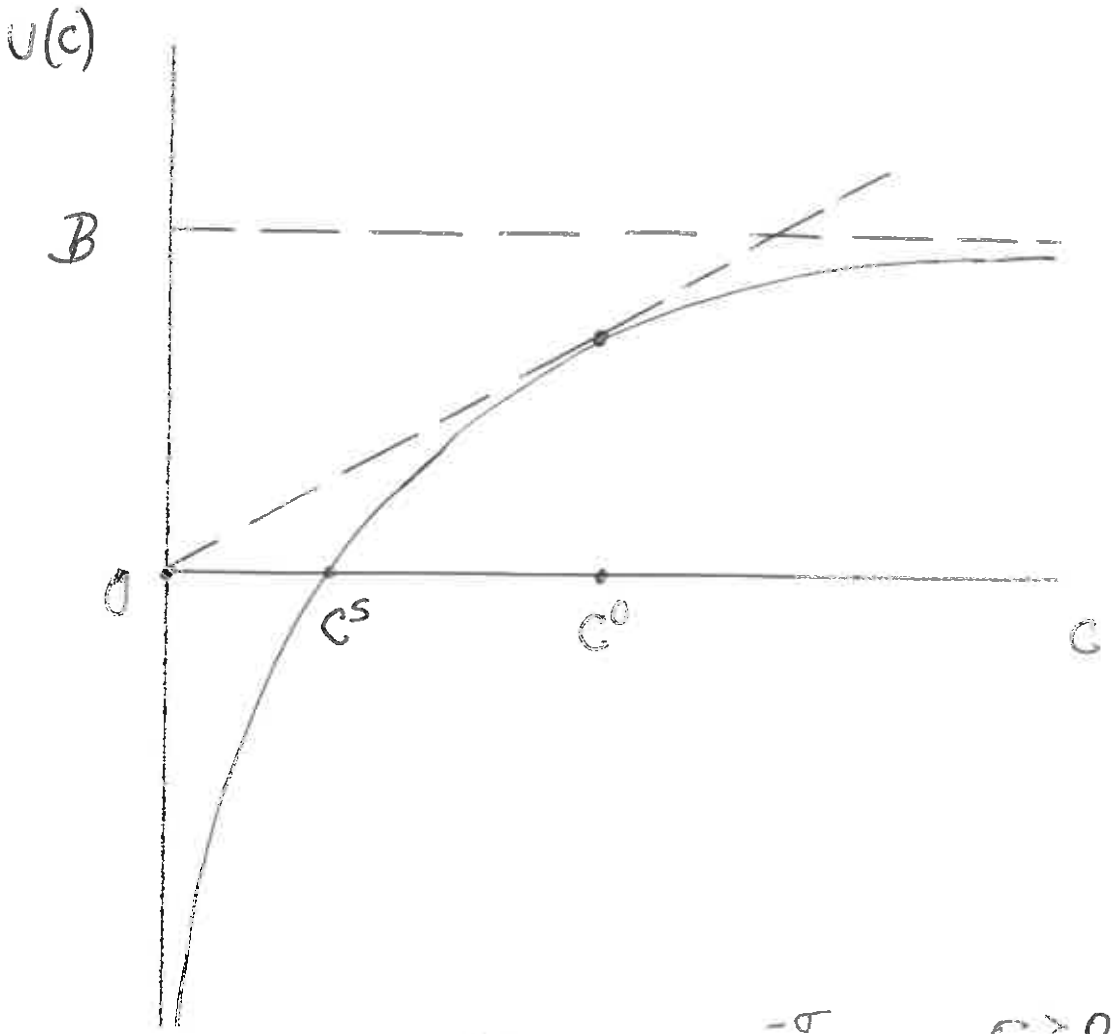


Figure 2



$$U(c) = B - c^{-\sigma} \quad , \quad \sigma > 0.$$

Figure 3

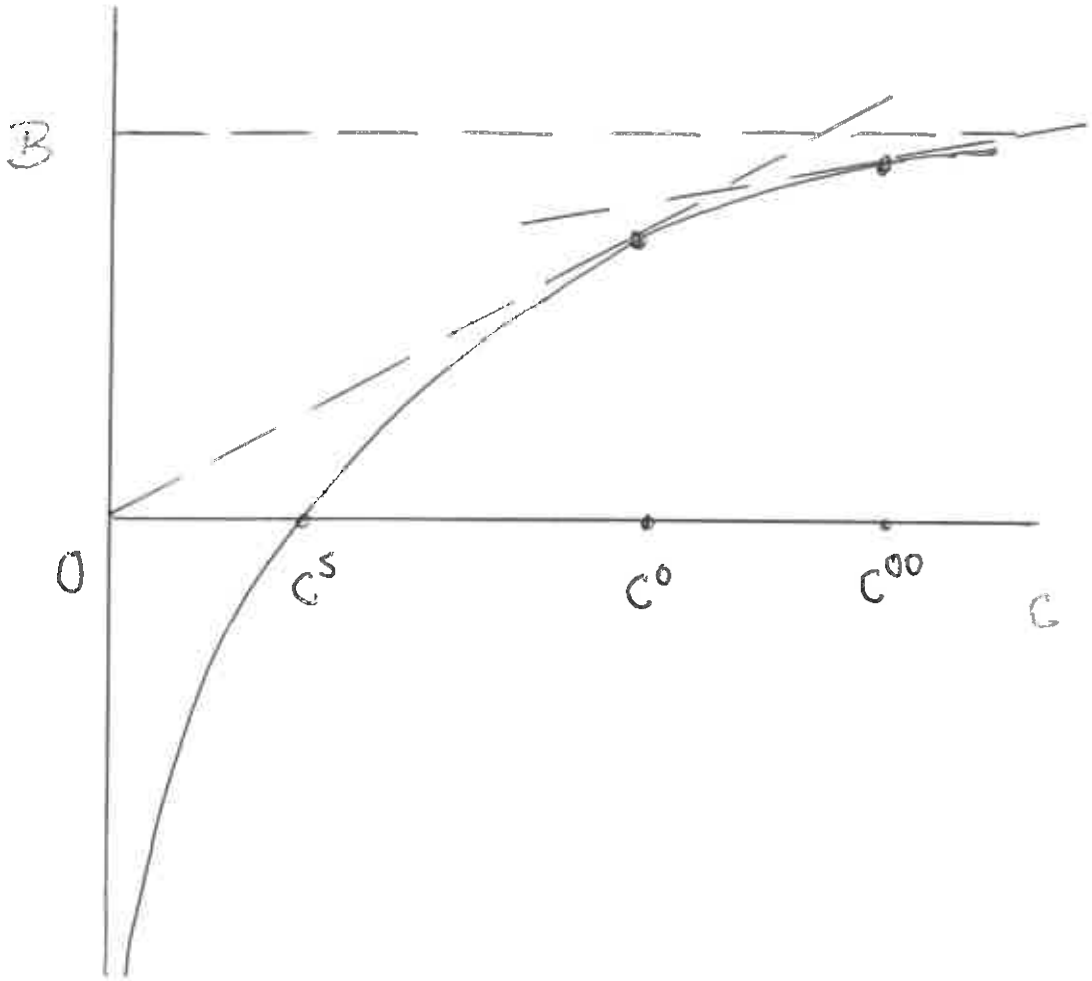


figure 4