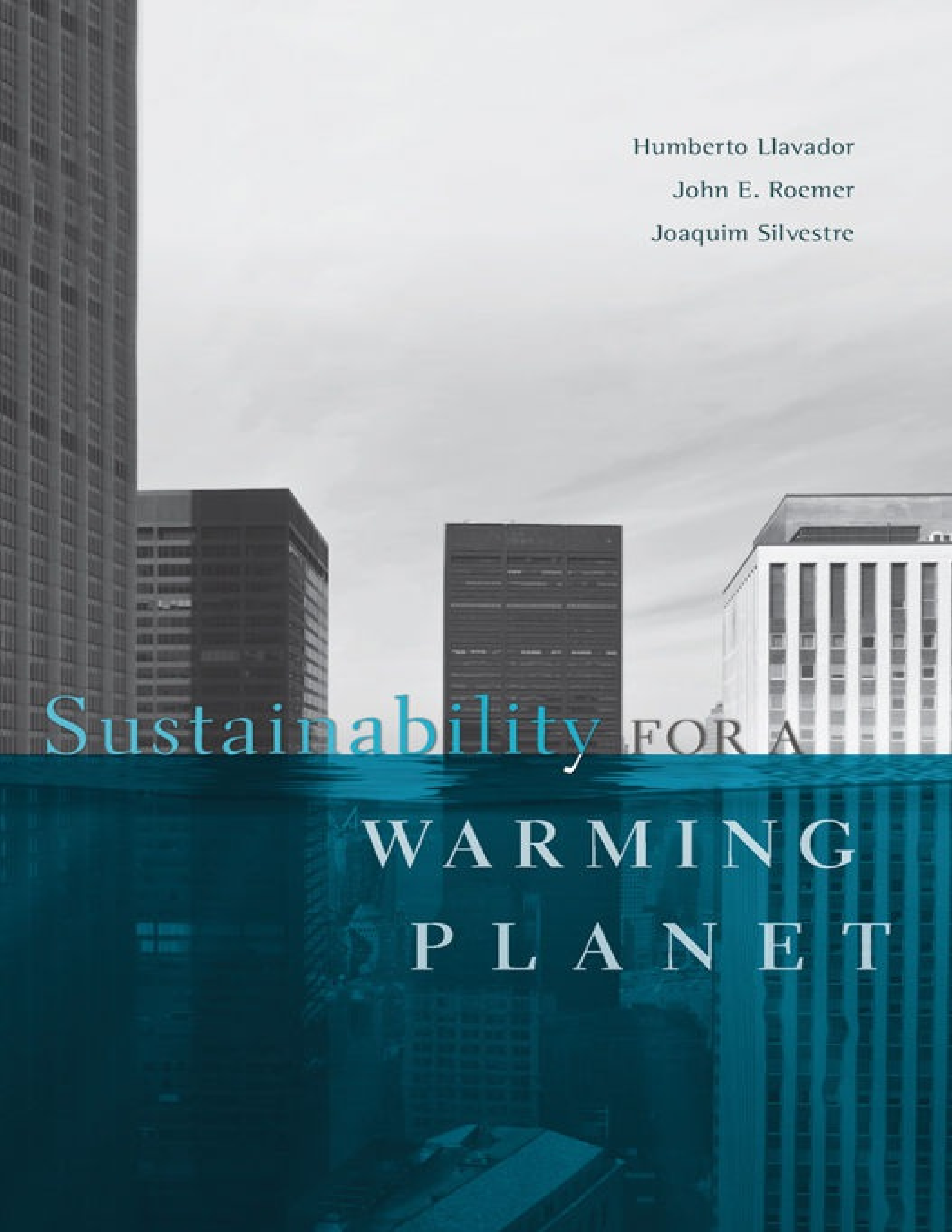
Humberto Llavador
John E. Roemer
Joaquim Silvestre
Sustainability FOR A
WARMING
PLANET

SUSTAINABILITY

for a WARMING PLANET

Humberto **Llavador**

John E. **Roemer**

Joaquim **Silvestre**

Harvard University Press

Cambridge, Massachusetts
London, England
2015

—-1
—0
—+1

Contents

—-1
—0
—+1

Introduction

Perhaps the single word that best summarizes the ethos of those who are concerned with climate change and its effects, instigated by carbon emissions since the advent of the Industrial Revolution, is sustainability. It is therefore remarkable that sustainability appears to be of tangential concern in some of the most prominent work of economists who have concerned themselves with this 'greatest of all externalities,' in the words of the *Stern Review* (Stern 2007), that is, with man-made global warming. The word does not appear in the index of either *A Question of Balance* (2008) or *The Climate Casino* (2013), the most recent books by William Nordhaus, one of the most prominent economists who studies global warming. *The Economics of Climate Change: The Stern Review* does refer to sustainability, but this is perhaps an exception. There is an older tradition in economic theory concerned with sustainability: notably, Solow (1974, 1993) proposed what can be thought of as the intellectual ancestor of this book. We will try to explain this disjuncture between the common (although not universal) practice of economists and popular expression and concern.

Solow (1993, 168) proposed a conception of sustainability and linked it to an ethical view. He wrote, 'I will assume that a sustainable path for the national economy is one that allows every future generation the option of being as well off as its predecessors. The duty imposed by sustainability is to bequeath to posterity not any particular thing . . . but rather to endow them with whatever it takes to achieve a standard of living at least as good as our own and to look after their next generation similarly.'

Two points in this quotation are noteworthy: first, Solow's concern is with sustaining a standard of living, not particular physical or natural assets; second, the conception is unidirectional in time, in that it is said that we must guarantee to future generations at least as good a living standard as we have. We follow Solow regarding the first point, but we extend his unidirectionality with respect to time into bidirectionality. We add that we, the present generation, also have a *right* to be as well off as future generations. Does this view preclude economic growth? We will argue that it does not.

Our attempt, in this book, is to bridge the gap between the concerns of billions of people to put our species' behavior on a sustainable path and economic theory. We propose a formal definition and economic model of sustainability and compute what sustainable paths of economic activity look like. We attempt to answer the questions: Is it possible to sustain human life indefinitely at a *decent* quality level, which means, among other things, limiting carbon emissions so as to control global temperature and climate? Is it possible to do so and to allow the developing world to reach the income levels of the rich world? If so, how would we have to change our economic practices in order to carry out an optimal plan? How much growth would have to be sacrificed?

We define several versions of sustainability in our study. None of them, perhaps, is as radical as the version many environmentalists would advocate—what's called strong sustainability, which means to preserve the flora and fauna of the earth, just as they are, or would be, without human interference. Our approach is anthropocentric and follows Solow: we define a *purely sustainable practice* as one that will maintain human welfare, of all generations now and into the indefinite future, at the highest possible base level or greater, as far as we can tell. We attempt to compute what that base level would be, on the path that maximizes it, subject to many assumptions about how the economy works—in particular, how technological change will respond to investments we make in knowledge creation and how the skills of workers will respond to education.

Pure sustainability, as some readers will recognize, can be phrased in a way that is familiar to those who are acquainted with John Rawls's theory of justice. It requires us to find the path that maximizes the welfare level that can be enjoyed by all generations into the future. In particular, it maximizes the welfare of the least well-off generation. (We abstract, initially, from differences of welfare *within* a generation of people and consider only the average welfare level of each generation in order to focus attention on the problem of *intergenerational equity*. Later, in Chapter 5, we introduce a concern for welfare differences across regions of the world at each generation.) It will transpire that the optimal sustainable path—the one that *guarantees* to all generations who live the highest possible level of welfare—actually *equalizes* the levels of welfare that each generation will experience. No generation would have a welfare level higher than any other in the problem's solution! There would be, on this path, no economic growth, where that growth is conceived of as a growth in welfare over time.

This fact may strike some—or many—as undesirable, perhaps sufficiently so to recommend scuttling this approach to sustainability. It is therefore important to explain why this 'no growth' result occurs. In our model of economic activity, there are a number of ways (four, to be precise) that each generation passes down important resources to the next generation: the education of its children; investment in research and knowledge, which lasts into the next generation; preservation of the biosphere; and investment to replace depreciated capital stock and perhaps augment it. These four activities all require sacrifices by early generations for later ones. Let us imagine, to keep the story simple, that there are only two generations: the first generation contains adults and children, and the second generation contains just adults who matured from those children. Suppose we want to maximize the level of welfare that both generations (of adults) can enjoy, and suppose the solution to the problem were one in which the welfare of Generation 2, u_2, were greater than u_1. So, the largest level of welfare that could be guaranteed to both generations is, by hypothesis, u_1. Now this can't be the correct solution: for have the first generation pass down fewer resources (say, by educating its youth a little less) and keep those resources for itself (adults in Generation 1 teach less and instead manufacture more consumer goods for themselves). This will increase u_1 a little bit and reduce u_2 a little bit, showing that the original solution did not, indeed, maximize the level of welfare that could be guaranteed to all generations. It is this argument, in the much more complicated model that we construct, that shows that on the optimal pure-sustainable path, all generations must enjoy the same level of welfare.

Indeed, we say that each generation has a *right* to insist that this 'maximin' path be chosen. Why? Because we view the date at which a person is born to be a matter of luck—a circumstance beyond his or her control—and we think it is *wrong* that persons be forced to bear the consequences of unchosen luck, if that can be avoided. John Rawls gave exactly this motivation for supporting his 'difference principle': that, in his words, circumstances that are *morally arbitrary* should not affect the conditions of persons. In Rawlsian parlance, we would say that the date at which a person is born is morally arbitrary, and this means that concerns of intergenerational equity require that human society set itself the task of guaranteeing to all who ever live the highest possible level of welfare that can be so guaranteed. At least, this is the *first step* in our argument.

There is, however, a catch. Persons do not have to enforce their rights. They may choose not to enforce this particular right if, for example, they would *prefer* that those who populate future generations be better off than they. We believe that this preference is, indeed, held by most people. Not only would most of us like our children to be better off than we are, but that kind of altruism toward future beings extends beyond our own line of descendants. Perhaps this is because we value human development as a good: we see how much better our lives are than those who lived in the past, and we are willing to sacrifice, up to a point, so that those who live in the future can be better off than we are. Perhaps we desire that humans accomplish great things—learn about the origin of the universe, find a cure for cancer—and these discoveries can only happen if we educate future generations more than we were educated. We do not attempt to test this conjecture but assume that it is so. This motivates our second definition of sustainability, which is to *sustain growth.* Growth sustainability (say, at 25% per generation) means to find that path of economic activity that maximizes the welfare of the present generation, subject to guaranteeing that welfare grows at least at 25% per generation, forever after.

But how do we find the right growth rate? Is it 25% or 50%? According to what we've said, this depends on how much each generation *would like* to sacrifice to enable future humans to enjoy greater welfare than they. We do not try to answer this empirical question, but what we do is compute paths of sustainable growth for several (we think reasonable) growth rates. For instance, a growth rate of welfare of 1% per annum means a growth rate of welfare between generations, each of which is postulated to be active for twenty-five years in our model, of 28.2%, while a growth rate of 2% per annum engenders a generational growth rate of 64.1%.

Pure sustainability and growth sustainability are the main, but not the only, concepts that we employ in this book. Suppose that we calculate the optimal path that sustains a 28.2% generational growth in welfare, forever. Clearly, the first generation will be worse off than it would have been on the pure-sustainable path (where the growth rate of welfare is 0% per generation) because it must invest a little more and consume a little less than it would on the pure-sustainable path. How much worse off would its members be? Very little, it turns out: the sacrifice in welfare that the first generation must sustain in order to potentiate welfare growth of 28.2% per generation forever is less than 1% of its welfare on the pure-sustainable path (see Chapter 3). Indeed, all generations except the first one turn out to be better off on the 28.2% growth path than on the pure-sustainable path. (In other words, only the first generation has to 'sacrifice,' compared to the benchmark of pure sustainability, to guarantee 1% per annum growth indefinitely.) However, at each date the current generation would be better off still if it were to enforce its right to insist on the pure-sustainable path *from that date onward*. We are, however, assuming that *each* generation desires to render future generations better off than they and so the growth path will be voluntarily sustained indefinitely.

How do we measure welfare? We take human welfare, or utility, to be a function of four inputs: material consumption; leisure time multiplied by a factor reflecting one's educational level; the stock of human knowledge, to date; and the quality of the biosphere, which is postulated to decrease as the atmospheric carbon concentration increases. The first two inputs are private goods, and the last two are public goods. We break with economic tradition in two ways: in valuing a person's leisure time by his or her educational level and in including the stock of human knowledge in the utility function. We believe that education increases the diversity of uses of leisure time, and the ability to do many different things with one's leisure increases welfare. We resist the neoclassical tradition in economic theory of viewing education as simply instrumental, as a means to increase future earning power. Likewise, neoclassical economics treats knowledge creation as an activity that is undertaken by a society to increase material wealth and welfare. But this is clearly not the entire story: much of our social investment in the creation of new knowledge (as opposed to educating young people about existing knowledge) is motivated by our *desire to know*. (Think of the $5 billion Hadron collider at the European Organization for Nuclear Research [CERN]. It would be provincial to attempt to justify this investment only on the basis of future inventions that small-particle physics will engender.) This applies not only to the arts and literature, which clearly elevate the human condition, but to mathematics and science as well.

Our utility function, however, does not include a concern for future generations—which, as we've said, we also believe people have. So, it's not intended to capture everything about what's valuable to people but rather what we might call their personal condition, their self-interested welfare. It is a version of the standard of living to which Solow refers. We attempt to address the altruism that people direct to future generations through the concept of growth sustainability, as we explained above.

What we've described so far is how we define sustainability in a model with certainty. However, uncertainty is an essential part of the climate-change problem; we will describe how we modify the model to include uncertainty.

Having explained, briefly, our ethics and the economic approach it leads to, we now briefly explain what most economists who study climate change do. They do not study paths of economic activity that sustain human welfare in either of the two senses that we have defined. Instead, they compute paths that maximize a weighted *sum* of generational utilities, into the distant future, where the weight put on the utility of each future generation declines geometrically (i.e., by a fixed factor at each generation). This is called a *discounted-utilitarian objective*.

There are a number of possible justifications for being interested in the path of economic activity that maximizes a discounted sum of generational utilities, as we explain in Chapter 1. The main ones are the following: Until John Rawls's work, utilitarianism—the view that justice consists in maximizing the sum of utilities of all people in society—was the ruling philosophical view, developed by Jeremy Bentham, John Stuart Mill, and

Henry Sidgwick, among other philosophers in the nineteenth century. Applied to a society that lives over many generations, utilitarianism would say to maximize the sum of utilities of all generations. Let us, for the moment, assume that the horizon is finite—we are concerned with the next five hundred generations only—and that the population is of unchanging size over time. Then the unvarnished utilitarian would seek the path of resource allocation that maximizes the total utility that these five hundred generations enjoy. The major criticism of utilitarianism, leveled not only by Rawls but many others, is that it is indifferent to interpersonal (in this case, intergenerational) inequality. Its sole concern is to maximize the *total* utility experienced, and the allocation of that utility among the generations is entirely subservient to this end. With a two-generation society, a utility profile of (99, 1) would be preferred to a profile of (49, 49), because 100 is bigger than 98.

One might be tempted to respond to this example with the claim that humans exhibit diminishing marginal utility (in resource consumption). But by hypothesis, the numbers given in the above example are *utilities*, not *resource levels. Whatever* resource levels give rise to these two possible utility distributions, as long as they are both feasible, are irrelevant: the utilitarian must prefer a utility distribution of (99, 1) to one of (49, 49).

One might also be tempted to respond to this example by saying that each generation should be assumed to have the same utility function (i.e., the same technique for converting resources into welfare). A 'utility monster' is an individual who converts resources into utility especially effectively and so a utilitarian must assign to her the lion's share of resources. However, even if we assume that every generation possesses the same utility function, the problem of the utility monster can arise easily in an intertemporal context because of technological change. Let's assume that technological change occurs over time, as indeed it has since the Industrial Revolution. Then to maximize total utility over five hundred generations, it might be that we should devote a lot of resources to technological innovation in the early generations, leaving these folks very little to consume, so that later generations have a wonderful level of technology, which can convert resources very efficiently into human welfare. Even though every generation may have the same utility function over consumption goods, resources will be able to produce consumption goods in much greater abundance in the future, because of technological innovation, than at present. Technological change makes later generations into utility monsters.[1]

To deal with this problem (and another one, discussed next), one can *discount* the utility of future generations—that is, make it count for less, the farther out in time it is experienced. It should be clear this is an *ad hoc* solution to the problem. In particular, one would have to tailor the discount factor to the rate of technological change. There is no clear, general, ethical instruction as to what it should be. Of course, if technological change is endogenous (as it is in reality—i.e., the investment in research and development is part of the optimization problem), the choice of discount rate is even murkier. The problem is with utilitarianism, and that is addressed by recognizing it as a poor ethical doctrine, especially in situations in which utility monsters can occur, as with intertemporal optimization and technological change.

There is a second problem with utilitarianism in intertemporal economic analysis. Now, however, we assume an infinite time horizon: although humans will eventually disappear, we do not know when this will occur; therefore, there is a strong argument for postulating a possibly infinite time horizon. (We do this throughout the book.) Again, technological progress generates a problem but now of a different sort. It turns out that in most economic models there are many paths (which extend into the infinite future) upon which total utility is *infinite.* Utilitarianism does not tell us how to choose among these. There are two ways of escaping this problem. The first way recognizes that the human species will not exist forever. Suppose we model this by saying that there is a fixed probability π that each generation, conditional upon its coming into being, will be the last one. We call this probability the (*generational*) *hazard rate*. If there is a utilitarian social planner or an Ethical Observer who wishes to maximize the sum of utilities of all those generations that, in the event, exist, subject to the stochastic process described, it turns out that it (or he or she) should maximize the discounted sum of generational utilities (over an infinite

[1] It is not that the individual who converts resources very efficiently into welfare is a monster *per se*: it is that following a utilitarian ethic converts this person into a monster who monopolizes the resource endowment.

future span) where the utility of Generation t is discounted (or multiplied) by the factor $(1-\pi)^{t-1}$. (See Section 1.2.2, third justification.) This is the first escape from the problem of infinite sums. If the factor $1-\pi$ is small enough, then the problem of maximizing the infinite *discounted* sum of utilities, over time, will have a unique, finite solution.

The second way of evading the problem of divergent total utility is to make an analogy with *individual* utility. It is not unreasonable to model a *person's* utility as the discounted sum of his annual utilities over his lifetime. Why might a person discount his future utilities? Because of impatience. From *today's viewpoint*, a 'util' today is more valuable than a util experienced in ten years' time—perhaps because of impatience or a 'defective telescopic faculty' or perhaps, entirely legitimately, because there is a chance the individual may not be around in ten years and so he rationally discounts that future utility. Now some economists move—we think inexcusably—from the reasonable model of a person as someone who maximizes the discounted sum of his or her utilities over periods of life to the unreasonable conclusion that society should maximize the discounted sum of utilities of future generations—as if those generations were simply phases of the life of one infinitely lived being. This is the second way of 'justifying' an intergenerational ethic that maximizes the discounted sum of generational utilities.

While the first solution to the problem of the infinitude of the utilitarian sum descends from the philosophical view of utilitarianism, augmented with the uncertainty represented by the hazard rate, the second derives from an analogy between a human species living for many generations and an 'infinitely lived consumer.' The first solution is the one taken in the *Stern Review;* the second is, essentially, that one taken by William Nordhaus. Stern and Nordhaus have sharp differences in approach, which follow from the different justifications they employ for the 'discounted-utilitarian' approach. We discuss these in detail in Chapters 1 and 4. Essentially, in terms of the mathematics, their different justifications lead to employing different discount factors to apply to the utility of future generations. Stern's discount factor is a version of the '$1-\pi$' that we have described above, and Nordhaus calibrates his discount factor to the rate of impatience of present-day consumers, as estimated from observed market interest rates. Stern's discount factor—the number by which he multiplies the utility of the *t*th generation—is larger than Nordhaus's, which means he gives more weight to future generations in his objective function than does Nordhaus.

We, however, disagree with both of these justifications for taking the objective to be a discounted-utilitarian one: with Stern because we are not utilitarians (we are *sustainabilitarians*) and with Nordhaus because we reject the analogy between an impatient infinitely lived consumer and an indefinitely long sequence of generations of a human society. In particular, Stern's approach is still vulnerable to the 'utility monster' problem because he chooses a discount factor close to one. We present our opposition to utilitarianism as a good theory of intergenerational justice in Chapter 1.

It may seem surprising that *so many* economists who work on climate change use the discounted-utilitarian model: and indeed, probably most of them follow Nordhaus's lead rather than Stern's. We give citations in Chapter 1 that demonstrate that the justifications for doing so are as lacking in philosophical substance as we have just indicated.

In sum, there are essentially three reasons for adopting a discounted sum of generational utilities as the social objective function:

(1) Because future generations will be better off than we are, we should penalize them by discounting their utility in the social objective.
(2) If the Ethical Observer is a utilitarian but understands that the number of generations that exist is a stochastic event, then she should discount future utilities by a discount factor related to the probability of extinction.
(3) Since many consumers discount their own future utilities due to impatience, it is therefore rational that society should discount the utilities of those who exist later in time.

We reject (1) as entirely *ad hoc* and (3) because the analogy between the consumer who lives many years and a society of different generations does not hold water. We find (2) to be philosophically consistent, but we reject it because we reject utilitarianism as a philosophy due to its indifference toward inequality.

As sustainabilitarians, we also recognize the necessity of discounting because of the uncertainty of the human species' lifetime. But discounting with sustainabilitarianism leads to very different results than with utilitarianism, as the reader will learn.

The core of our analysis, presented in Chapter 3 (the world as a single region) and Chapter 5 (two regions, North and South), involves theory and calibration. It postulates the four above-mentioned intergenerational links (education, knowledge, physical capital, and the quality of the biosphere) and includes the education level and the state of knowledge in a person's quality of life, welfare, or utility function. Our computational approach is inspired by turnpike theory. We prove a turnpike theorem in the simpler, theoretical model of Chapter 2, which has a preliminary character and establishes a link with the familiar Ramsey model, widely used in growth theory. The first-order condition of the Ramsey model (see the Chapter 1 Appendix) yields the Ramsey equation, frequently used by climate-change economists as a calibration tool (Chapter 4).

The traditional Ramsey model has only one intergenerational link (physical capital), whereas our Chapter 2 model adds education. It contains two sectors of production: one sector, to be thought of as manufacturing, produces the only physical 'commodity,' which is used for consumption and investment, and the other sector is education. Manufacturing uses inputs of capital and skilled labor to produce its output, and the education sector uses only the labor of teachers as an input, to produce educated children, who become the next generation's skilled adult workers. As is customary in growth theory, we abstract away from intragenerational issues in that we assume that each generation consists of a single household (one adult, one child). The society's problem, at each date, is to partition the adult's skilled labor endowment among three uses—manufacturing, teaching, and leisure—and to allocate the output of manufacturing between adult consumption and capital investment. (One may either view children as not consuming at all or view the adult's consumption as household consumption. The child's utility, however, is not modeled.) Adults derive utility from two sources: consumption and their educated leisure time. Some may argue that it's a bit inconsistent to value leisure time by one's educational level, since there is no opera or soccer to choose between in this simple world, but all models take shortcuts.

The economy begins, at date zero, with an endowment consisting of adult skilled-teaching labor and a physical capital stock: call these two values x_0^e (for the stock of labor at date zero in the education sector) and S_0^k (for the stock of physical capital at date zero). Given this endowment vector, there is a set of *feasible paths* that specify all the details of economic activity in every generation, forever. Such a path specifies how each generation's adult allocates her time among the three activities (manufacturing, teaching, and leisure) and how it allocates the produced good between immediate consumption and investment to augment the capital stock. Each generation passes on to the next a stock of capital and education, embodied in its (young) adults. In this simple model, there is no knowledge creation, no technological change, no emissions, and no climate change. Nevertheless, the insights we garner from it will apply to the central models, with climate-relevant parts, of Chapters 3 and 5.

Among these paths, the problem is to choose the best sustainable one. In the pure-sustainable variant, optimization means to find the path that maximizes the highest utility that can be guaranteed to every generation. In the growth-sustainable variant, one finds the path that maximizes the first generation's utility, subject to guaranteeing some exogenously specified growth rate of utility per generation, forever.

As we said, in the pure-sustainable variant, on the optimal path, utility is held to a constant for all generations—there is no growth. The first major result is what is called a turnpike theorem, which makes two assertions. First: if the endowment vector (x_0^e, S_0^k) lies on a certain ray in the (x^e, S^k) plane, then the optimal solution to the pure-sustainability problem is a *stationary state*—that is, all economic variables are constant over time. The skill of the labor force is exactly reproduced at every generation, capital is maintained at exactly S_0^k forever, and consumption and the allocation of labor are unchanging over all time. Second: if the endowment vector is not on this special ray, then the optimal solution converges to one of the stationary states. Because, typically, convergence occurs quite rapidly, the turnpike theorem tells us that we get a good picture of what the optimal path of economic activity looks like by understanding what the stationary states look like.[2]

A turnpike theorem also holds for the growth-sustainable variant. There is a ray (which depends on the exogenous growth rate) such that, if the date-zero endowment vector lies on it, then the optimal solution to the problem

[2] Figure 2.1 in Chapter 2 provides an illustration of the turnpike theorem.

is a *balanced growth path*, which means that every economic variable grows at the specified growth rate, forever. If the date-zero endowment vector does not lie on this ray, then the optimal path converges to one of the balanced growth paths.

Let us observe that we can conceptually separate the set of *feasible paths of economic activity* from the *objective function* of the optimization program. For example, the set of feasible paths is the same for both the pure- and growth-sustainable models. It's just that the *optimal* path depends upon the particular objective function chosen. In particular, we can also maximize a discounted-utilitarian objective on the same set of feasible paths of our simple model. We characterize how small the discount factor must be in order for *that* program to have a solution. There turns out to be an exact formula for computing this discount factor, which depends (it so happens) on how productive the educational technology is. We can think of the productivity of the educational technology as being the inverse of the number of hours per year an adult has to spend in educating her child so that the child acquires the same level of skill as she. (There is no genetic variation here between adults and children.) Equivalently, one can think of educational productivity as the teacher-student ratio in a society in which teachers exactly reproduce the skill level of the adult generation in their children. This productivity is represented by the parameter ξ in the model, which, for the US economy, has a value of about 41.[3] It turns out that the discounted-utilitarian program in the simple model converges (i.e., possesses a solution and does not diverge to infinity) exactly when the discount factor is less than $1/\xi$. Suppose we take the discount factor to be $1-\pi$, where π is the per-generation 'hazard rate' of species extinction. Then the discounted-utilitarian objective will diverge when $1-\pi > \frac{1}{\xi} \simeq 0.024$, which means when $\pi < 0.976$. Of course, any reasonable generational hazard rate is far less than 97.6%, and the discounted-utilitarian objective diverges for the model of Chapter 2, given these parameter values.

[3] Roughly speaking, today's cohort of teachers can reproduce their own level of education or skill in today's generation of children with an average class size of forty-one. If class sizes are on average smaller (as they are), then today's teacher cohort will produce children whose average level of skill is greater than that of the adult cohort.

For a while (meaning until Chapter 6), we treat the probability of catastrophe as fixed—some small number, denoted π, that each generation will be the final one, assuming that the species has not already disappeared. This approach to uncertainty is taken, as well, in the *Stern Review*: indeed, Stern takes the catastrophic probability to be 0.1% per annum, which compounds to 2.47% for each twenty-five-year generation ($\pi = 0.0247$). We believe Stern's choice of per annum hazard rate is far too large. The human species has been around for about 100,000 years, or four thousand generations. Suppose the hazard rate has been constant over this period (perhaps until recently). What would that hazard rate have to have been in order that our species should have survived until now with a probability of one-half? It is easy to compute that the generational hazard rate would have been $0.00017 = 0.017\%$![4] We believe Stern's hazard rate is too large by two orders of magnitude ($0.0247/.00017 = 145$). If Stern's generational hazard rate of 2.47% were correct, the probability that the human species would have survived four thousand generations is 3.57×10^{-44}—we would not be here today.

Taking the catastrophic probability as exogenous and fixed is a shortcut to a more satisfactory approach, which would model the probability of catastrophe as a function of global emissions that raise carbon concentration, which has climatological consequences that increase the probability that the species will disappear. But endogenizing that probability, as just described, renders the analysis more subtle and will not be attempted until Chapter 6.

Given this uncertainty, with a constant, exogenous hazard rate, what is the appropriate objective to maximize? For a utilitarian or a sustainabilitarian? We return to the Ethical Observer, who represents the intergenerational ethical view. Suppose she is a utilitarian. Then, we say, she wants to maximize the sum of utility levels of all generations that, in the event, exist. The probability that any generation, T, will be the last one, is $\pi(1-\pi)^{T-1}$. Thus, if an infinite path of economic activities is chosen that generates the infinite path of utilities $(u_1, u_2, \ldots)$, she calculates the *expected value* of her objective function, which is: π times u_1 plus $\pi(1-\pi)$ times $(u_1 + u_2)$ plus

[4] If $(1-\pi)^{4000} = 0.5$, then $\pi = 0.00017$.

$\pi(1-\pi)^2$ times $(u_1+u_2+u_3)$ and so on, forever. The generic element in this sum is the probability that Generation T is the last one multiplied by the total utility experienced in the world in those T generations. This sum can be shown (Section 1.2.2) to equal the usual *discounted-utilitarian sum*, with a discount factor of $1-\pi$: that is, it reduces to the sum $\sum_{t=1}^{\infty}\pi(1-\pi)^{t-1}u_t$. So, as we've noted, for the model of Chapter 2, the discounted utilitarian-objective diverges for any reasonable value of π.

Now suppose the Ethical Observer is a sustainabilitarian. The same reasoning shows that her expected value, on a path, is equal to π times u_1 plus $\pi(1-\pi)$ times the minimum of u_1 and u_2 plus $\pi(1-\pi)^2$ times the minimum of u_1, u_2, and u_3, and so on, forever.

The computation of the optimal solution to the program of the discounting sustainabilitarian might be a difficult undertaking, but we prove a 'simplification theorem,' which asserts the following. If π is sufficiently small, then the solution to the *undiscounted* pure-sustainabilitarian problem is identical to the solution of the *discounted* pure-sustainabilitarian problem. How small does π have to be? Exactly the value for which the *discounted-utilitarian* program just diverges with the discount factor $1-\pi$! Therefore, because we have characterized the exact value of the discount factor that causes the discounted-utilitarian program to diverge, we know how small π has to be for the premise of the simplification theorem to hold. In particular, that premise holds as long as $\pi < 0.976$.

The simplification theorem says that, if π is small enough, then the sustainabilitarian *does not have to discount*. This doesn't mean that *conceptually* she does not have to discount—indeed she does, because future generations might not exist, and this must be taken into account. It says, rather, that *computationally*, discounting is not necessary. Discounting turns out not to matter for the result. This is very different from the situation for the utilitarian Ethical Observer, in which discounting is vital—for the optimal solution to the discounted-utilitarian program depends upon the exact value of the discount factor. We attempt to give, in Chapter 2, some intuition for why the simplification theorem holds. The key lies in the high productivity of the education sector—that the adult generation need devote only a very small fraction of its time to reproduce, in its children, the same level of skill as it possesses.

Chapter 3 builds, calibrates, and solves our full intergenerational model, with emissions, climate change, and technological innovation. The economy now has three sectors: manufacturing and education, as in the simple model, and in addition, a knowledge sector, which employs workers to produce knowledge. Knowledge is conceived of as art, literature, science, mathematics, technological innovation, and so on. It is produced, in our actual economy, mainly by researchers in universities and research institutes and in the research divisions of firms. The knowledge and education sectors are different: educators implant existing knowledge in the heads of students, and researchers discover new knowledge but do not teach. (In calibrating the model to actual economies, we must allocate the time of university professors between their research and teaching functions.)

In addition, we now introduce carbon emissions, which are produced by the manufacturing sector. These emissions increase the concentration of carbon in the atmosphere. The concentration of carbon is a public bad, which enters into the economy in two ways. First, it directly diminishes human welfare and second, it enters negatively into the production function of the manufactured commodity. It is harder to produce goods when climate events occur.[5]

Figure I.1 provides a summary of how the economy works. Human welfare now results from five inputs: consumption of commodities, leisure valued at its level of education, the stock of human knowledge, and biospheric quality, which is represented (negatively) by the stock of atmospheric carbon. We have explained above why we consider knowledge a generator of human welfare. There are, as we said, three production sectors. Commodity production now uses as inputs skilled labor, capital, and knowledge, and it is adversely affected by the public bad of carbon concentration. Production produces both emissions and output, an aggregate of all commodities.[6] The knowledge input accounts for technological progress in commodity production. There are now four intergenerational transfers: a stock of capital,

[5] In particular, this may be so for agriculture, although some argue that a moderate increase in global temperature will increase global agricultural production.

[6] Formally, emissions are modeled as a production *input*. The more emissions are allowed, the more of the commodity can be produced, so mathematically, emissions can be viewed as an input in production. Compare with Section 4.3.2.

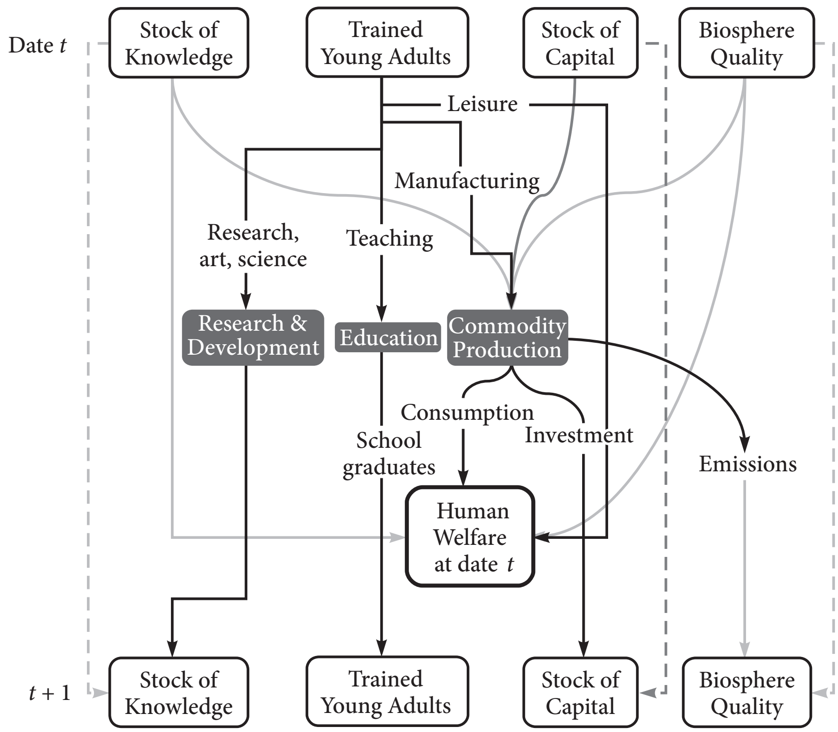

Figure I.1. A flowchart of economic activity from one generation to the next.

skill embodied in young adults, a stock of knowledge, and a stock of biospheric quality (the non-carbon-polluted atmosphere).

The economy begins with a vector of endowments with four components: a stock of capital, educated teachers, a stock of knowledge, and a level of biospheric quality (or, a stock of 'clean' atmosphere). Given these, one can generate the set of feasible paths of economic activity, and along each path, there will be an associated sequence of utilities for all generations.

As before, one can choose a variety of objective functions—in particular, those of pure sustainability, those of growth sustainability, or those of discounted utilitarianism.

We now assume that the main theoretical results concerning the simple model of Chapter 2 are true, as well, of the full models of Chapters 3 and 5. In particular, we assume that turnpike theorems hold. Indeed, we prove the first part of the turnpike theorem for the model of Chapter 3—namely, that if the endowment vector lies on a particular ray in four-space, then the optimal solution to the pure-sustainability program (or growth-sustainability program) is a stationary state (respectively, a balanced growth path). We *assume* that the second part of the turnpike theorem carries over—namely, that if the initial endowment vector does not lie on the ray, then the optimal solution to the program converges to a stationary state (or balanced growth path).

Chapter 3 calibrates the parameters of the model to an actual economy—namely, the US economy. It would be ideal to calibrate the model to the world economy; however, that would introduce large inaccuracies due to the unreliability or nonavailability of data. For the United States, we can be quite confident that the data are reliable. Calibrating the model means attaching numbers to the parameters in the three production functions, to the arguments of the utility function, and to how carbon emissions generate changes in atmospheric carbon concentration. Our calibration methods are explained in Appendix A: Calibration. They follow standard calibration practice and sometimes use data from the *Stern Review* and from Nordhaus (2013, 2008a, 2008b).

Having calibrated the model, we are ready to compute optimal solutions for the programs with the pure- and growth-sustainability objective functions. We take date zero to be 2010, so the endowment vector with which the economy starts is the vector of endowments of the US economy in 2010. The adult in the first generation of the model lives from 2011 to 2035.

A full specification of the optimization problem would require that we optimize over the path of carbon emissions as well as over all other economic variables. This would require estimating the functions that relate emissions to atmospheric carbon concentration, carbon concentration to global temperature, and global temperature to human welfare. The first of these functions is relatively well known; there is a great deal of debate about the second one, and about the third one, we can only make educated guesses. A full *integrated assessment model* (IAM) contains these three equations as well as all the equations we have specified. We have elected not to construct a full IAM because we believe the uncertainties in the specification of the second and third processes are so great. (However, in Chapter 6, we do construct a highly stylized integrated assessment model in order to make some ballpark calculations of how to respond to uncertainty.)

514-60040_ch01_1P.indd 18 1/5/15 9:56 PM

We adopt, instead, a *pis aller*. We constrain carbon concentration to follow a path that converges to a concentration of carbon below 450 parts per million (ppm) in seventy-five years (three generations of our model).[7] Since our model is of the US economy, we therefore constrain the emissions of the manufacturing sector in the model to be a certain fraction of emissions that are globally consistent with staying on the chosen path of carbon concentration. We compute two scenarios: one in which the United States continues to emit 17% of global emissions (its share of total emissions in 2010) and a second in which the United States emits roughly its per capita share of global emissions. Subject to each of these constraints, we solve the sustainability optimization programs (pure, and for various growth rates). Features of the optimal paths are presented in Chapter 3.

We will not rehearse the results here, as their verbal discussion in Section 3.11 provides a nontechnical summary. We do, however, indicate some of their main features. On the pure-sustainable path, human welfare, which is constant, is higher than it was in the base year (2010). Thus, following a sustainable path, which radically reduces carbon emissions over seventy-five years, is technologically feasible, without reduction in human welfare. Indeed, even growth is feasible, at least at moderate growth rates, while respecting the constraint imposed on carbon emissions. This is so for both scenarios—even when the US economy is restricted to emitting only *its per capita share* of global emissions! One might suspect that this is accomplished by radically changing our 'consumption bundle' away from commodities to enjoying more leisure with higher education: for recall, it is only commodity production in the model that generates greenhouse gas (GHG) emissions. This is not so: commodity consumption is in fact higher on the pure-sustainable path than it was in 2010 and becomes much higher on the growth-sustainable paths. And leisure time does not rise above its present levels on the optimal paths—in fact, it falls a little (although its value increases because educational levels increase). Indeed, this may be one of the surprises of addressing global warming: on the optimal path,

[7] The path is constructed from the Representative Concentration Pathway RCP2.6 and is described in Section 3.4.2. The RCP2.6 is representative of precautionary scenarios with very low GHG concentration levels and provides an expected temperature change below 2°C.

we do *not* continue to reduce the length of the working year or working life, as has occurred, dramatically, over the last century.[8]

Now there is, indeed, some change in the consumption bundle, but the main feature of the optimal paths is a sharply increased investment in the creation of knowledge: the fraction of labor engaged in the knowledge industry doubles compared to the benchmark year. Recall that the stock of knowledge is an input into human welfare. The other effect of this new knowledge is to induce technological progress that enables the economy to produce commodities with a much lower emissions/output ratio. A skeptic might question this part of our model: is it realistic to assume that knowledge investments will enable us to reduce the emissions/output ratio as sharply as the model predicts? We must answer by referring to our calibration of the production function, which uses the historical effect of knowledge on that ratio. Indeed, we might conjecture that in the future, research and development (R&D) will have an even stronger impact on the emissions/output ratio than it has in the last fifty years because R&D will focus on reducing that ratio. After all, the urgency of reducing emissions has only become apparent in recent years. Therefore, we do not view paths of feasible output and emissions that our model claims are feasible as reflecting a particularly rosy view.

There are several technical issues that the analysis in Chapter 3 addresses. Paramount among these is how to compute the optimal solution to the pure- and growth-sustainable programs. Consider, for example, the former optimization program. The problem exists because the vector of aggregate endowments at date zero (2010), with which the economy begins, does not lie on the ray from which the optimal solution is a stationary state, a ray that we can compute precisely and analytically. We know the optimal solution will converge to a stationary state, so the endowment vector will converge to a point on that ray, but which point? We solve this problem by an approximation technique that forces convergence in two generations. The technique gives us a very close approximation to the actual optimal path.

[8] In 2012 German workers labored 1400 hours per year on average, which amortizes to 28 hours per week for a 50-week year. In 1900 the average work week was 60 hours, or about 3000 hours per annum.

One of the main lessons of Chapter 3 is that, considering only the US economy, the problem of sustainability is solvable, as far as the laws of economics are concerned. It is pretty clear that there are no technological barriers to finding a sustainable path. The real issue is political: Will our society be able to agree on the changes needed to implement a sustainable path? Will we be able to muzzle the energy firms that see their profitability tied to the exploitation of fossil fuels? Will American consumers see the necessity of paying much higher fuel taxes in order to reduce the consumption of fossil fuels? In Chapter 5 we complicate the political problem by addressing the issue of global sustainability, but we do not, in this book, study the political and psychological issues that must be raised. Our task here is to present an economic and philosophical argument.

Section 3.9 carries out an experiment in which we ask: what would the optimal path look like if we chose to sustain *consumption* rather than *utility*? Both Stern and Nordhaus view utility as consisting only of commodity consumption. Some may view this as a useful simplifying move, but we have argued that it is important to include education, knowledge, and the quality of the biosphere in the generational utility function because this engenders the possibility of substituting from commodity consumption to these other goods in order to maintain welfare. In this section, we compare the pure-sustainability problem for *consumption only* with our optimal path for sustaining utility that we have computed.

The results are fairly dramatic. Unsurprisingly, the steady-state consumption when only consumption is sustained is somewhat higher than when utility is sustained: about 16% higher. However, both education and knowledge are much higher when utility is sustained: education is 385% times as much in the utility-sustaining program as in the consumption-sustaining program, and knowledge is 224% times as large. If we compute *utility* at the steady state of the consumption-maximizing program, it is only 44% of the utility at the steady state of the other program, achieving a mere 60.3% of the utility level in the 2010 reference year before the end of the century. In other words, a fairly small sacrifice in consumption enables a very large increase in education and knowledge creation. We believe these results underscore the importance of defining utility in an encompassing way.

Before introducing the issue of global cooperation to control carbon emissions, we proceed to Chapter 4, which looks more carefully at the approaches that economists William Nordhaus and Nicholas Stern have taken to climate-change analysis. Nordhaus's objective function appears to be an amalgam of a social-welfare function and a utility function for an infinitely lived consumer. Indeed, we present three interpretations of his objective function (Table 4.1) and argue that none of them is satisfactory, given our philosophical premise that the current generation has no right to impose its own time rate of discount upon the social objective. This is embodied in Nordhaus's choice of discount rate, of 0.015 per annum, which is calibrated from market interest rates, reflecting the degree of impatience of current savers and investors. Nordhaus's discount rate is associated with an annual *discount factor* of $\left(\frac{1}{1+0.015}\right)^{99} = 0.9852$. Consequently, he discounts the utility of people living one hundred years from now by a factor of $0.9852^{100} = 0.2256$; that is, their utility counts only 23% as much as today's generation's utility in the social calculus. We claim that this move cannot be justified by an appeal to an impartial Ethical Observer. One consequence of Nordhaus's parameterization of the model is that relatively mild reductions of GHG emissions are recommended in the short run; significant reductions only occur in the late part of this century. The projected temperature increases associated with Nordhaus's DICE 2013 model are considerably more than 2°C by 2100.

The *Stern Review* does not undertake a full optimization exercise, as Nordhaus (2013, 2008a, 2008b) does, but carries out a cost-benefit analysis for a small sample of emission paths. Basically, it argues that a particular alternative emissions path dominates the 'business-as-usual' (BAU) path in the sense that the benefits of adopting it are greater than the costs of doing so. An alternative path to the benchmark BAU path is recommended if and only if the present discounted value of differences in consumption between the two paths over the generations is positive. But what discount factor does Stern use? His discount factor equals $1-\pi = 0.999$ per annum, where the value $\pi = 0.001$ is his choice for the annual probability that a catastrophe will annihilate the human species. Thus, Stern discounts the utility of people one hundred years from now by a factor of $0.999^{100} = 0.9048$: their utility counts about 90% as much as the current generation's utility in the social calculus. As we've said, we find Stern's analysis to be philosophically consistent with a utilitarian theory of justice. It is, for that reason, impartial, unlike Nordhaus's analysis, but we nevertheless do not agree with Stern because we dissent from his utilitarianism. And as we've said, we find Stern's discount rate to be too large by two orders of magnitude (so his discount factor is *too small*).

We review, in Chapter 4, the criticisms that Nordhaus and Martin Weitzman have raised against the *Stern Review*. The best face we can put on Nordhaus's defense of his approach is that it would be politically a non-starter to propose discount rates of future generations' utilities that are far out of line with the discount rates that today's savers exhibit, as deduced from market interest rates. (It is the so-called Ramsey equation, which we discuss in Chapters 1 and 4, that links consumers' subjective discount rates to market interest rates.) The defense, in other words, is one of practical politics. Nordhaus writes, 'The time discount rate should be chosen along with the consumption elasticity so that the model *generates a path that resembles the actual real interest rate* [our italics—LRS].' Why must it resemble actual interest rates? Presumably because, did it not, it would be impossible, or extremely difficult, to induce consumers to alter their savings' behavior to fit the recommended path.

We believe climate-change economists must be clear on whether the analysis they are proposing is one based on ethical principles—that is, motivated by considerations of intergenerational (and intragenerational) justice—or one motivated by practical political considerations. There is surely need for both, but one must honestly distinguish between them. Stern's approach is clearly of the former type: it is based upon a utilitarian view of justice combined with a simple approach to uncertainty. As we've just said, Nordhaus's approach, of deducing the discount rate from actual, observed interest rates, which reflect rates of impatience of today's investors and savers, might be justified by a pragmatic, political aim. It might be pretty hard to induce citizens in a democracy to undertake behaviors that reflect very different attitudes toward future generations than their current behavior evidently exhibits. But Nordhaus does not make this argument. He maintains that his approach is purely objective and flows from sensible economic principles.

We believe that, at this point in the analysis, it is wiser to follow our ethics to see where they lead us. Our own approach, at least up until this point in the book, is unabashedly ethical. We do not claim that any rational person *must* be a sustainabilitarian or a growth sustainabilitarian.

We believe that many people will agree that they are one of these ethical types once they understand our definitions. But we cannot say it is crazy to be a utilitarian or to adopt some other conception of intergenerational justice.

We pursue the ethical approach because we believe that if a strong ethical argument can be made for limiting GHG emissions, that will in itself provide a strong rationale for changing patterns of behavior. The argument cannot be strengthened by building in constraints, at the beginning of the analysis, based on market behaviors that developed during a period when global warming was not an issue. In the end, compromises with ideal policies will be inevitable, based upon political realities, but the goalposts should be set prior to that point so that we know at what we should be aiming.

We conclude Chapter 4 by examining an attack that Nordhaus makes on those who, like Stern, apply much smaller discount rates than he. Nordhaus proposes a thought experiment in which scientists learn that a 'wrinkle' in the climate system will cause damages equal to 0.1% of net consumption per year starting in two hundred years and forever after. He points out that, with Stern's discount rate, we should be willing to pay 56% of one year's world consumption today to remove the wrinkle—about $30 trillion, approximately one-half of today's global gross domestic product (GDP), to fix this tiny problem in the distant future. As this strikes Nordhaus as absurd, it must follow that Stern's discount rate is far too low. We show that Nordhaus's example is misconceived, as it fails to anticipate that the proper response to the discovery of the climate wrinkle is an *optimal* response, not the poor response that Nordhaus proposes. We show that, *contra* Nordhaus, if the climate wrinkle were discovered, the *correct* response would be to reduce the consumption of the present and future generations up to when the climate wrinkle occurs by *something less than* 0.1%, in order to increase investments, which would reduce the damages to the people living two hundred years from now when the climate wrinkle occurs. Indeed, assuming that the original proposed path (before the wrinkle was discovered) was an *optimal* path, there is *no* path that will restore aggregate social welfare to what it would have been without the climate wrinkle.

Chapter 5 culminates our analysis by extending the model of Chapter 3 to a world with two regions. A new, *intragenerational* question arises: how should the burden of reducing GHG emissions be allocated between the developing world (the South) and the developed world (the North)? In other words, we add the problem of intragenerational welfare to the intergenerational issues studied in Chapter 3. To keep the analysis as simple as possible, we envisage a South that looks like China (i.e., that has the vector of aggregate endowments of present-day China) and a North that looks like the United States. Formally, the model now consists of an indefinitely long sequence of generations, where each generation consists of a set of identical households of the Chinese type and a set of identical households of the US type, where the numbers of these households are given by UN population estimates of the developing and developed world over the next fifty years. We assume that global population is unchanging after 2060.

Each region's economy is summarized by the relationships illustrated in Figure I.1, with two changes. First, it is assumed that knowledge diffuses from the region that possesses more knowledge to the region with less knowledge. The knowledge accruing to the South at Generation t is equal to the knowledge it has produced, plus the knowledge that has diffused to it from the North (assuming the North continues to be the region with a higher knowledge endowment). Second, each region can transfer commodities to the other. For example, it is conceivable that the North would specialize in the production of knowledge, the South would specialize in commodity production, and the South would export commodities to the North.

We exogenously adopt the same path of world emissions and concentrations as in Chapter 3, which, it will be recalled, provides an expected temperature increase above the preindustrial level of less than 2°C. The central question of Chapter 5 is: how should the global emissions allowance be allocated between the regions at each generation, and what are the optimal regional paths of production, consumption, education, and investment in knowledge?

We advocate policies that will lead to eventual convergence in the levels of welfare per capita in the various regions: this follows from our general egalitarian principle. But we do not have an ethical instruction concerning what the date of convergence should be. Here, we propose a political solution, that is, a particular solution to the bargaining problem that is taking place in the sequence of ongoing Conference of the Parties (COP) meetings that are concerned with resolving the issue of fairly sharing the burden of reducing GHG emissions.[9]

Many writers have proposed ethical solutions that are based upon a conception of the carbon budget that each nation has a right to claim. These claims are often deduced by assuming that, from some starting date in the past, each nation had a right to emit a total amount of carbon proportional to its population. If that date is taken, for example, to be 1990, then one would compute the amount of global GHG emissions that are acceptable from 1990 on, allocate those emissions to nations in proportion to their 1990 populations, and subtract from that endowment the emissions each nation has produced since 1990, to give the remaining budgets. Whatever the ethical merits of these proposals, we believe they are politically unrealistic, and so we search for a solution to which we think all parties could agree.

Our proposal is based on Thomas Schelling's focal-point approach to bargaining (Schelling 1960; Colman 2006). Often, Schelling says, there is a clear focal point in a bargaining problem, and this emerges as the agreement. We propose that there is such a focal point in the climate-change bargaining problem, and it concerns the date at which the developing world will catch up to the developed world in terms of welfare per capita or, more simply, GDP per capita. Our proposal is summarized as 'preserving the date of convergence.'

In our model, as we said, the world has two regions, a South like China and a North like the United States. Suppose that it were common knowledge that, absent the problem of climate change, the South would converge to the North in welfare or GDP per capita in seventy-five years. This is actually not a crazy estimate: we compute in Chapter 5 that, given a predictable decrease in China's rate of growth, a date of convergence of 2085 is a reasonable guess. Our claim is that any solution to the bargaining problem must preserve that date of convergence: preservation of that date is the focal point.

The argument is, as we said, political. Suppose the US negotiators were to propose an allocation of emissions' burdens whose consequence were

[9] The Conference of the Parties (COP) meetings are annual meetings of the parties in the United Nations Framework Convention on Climate Change (UNFCC).

that convergence would occur in one hundred years. The South would argue that that was unacceptable: why should the date of convergence be postponed for twenty-five years from what it would have been under BAU? Impartiality requires that the date of convergence be no farther than seventy-five years away. Conversely, suppose the Chinese negotiators were to propose an allocation of burdens whose consequence was convergence in fifty years: the North would never accept this, arguing that the date of convergence should not be advanced by twenty-five years. The only agreement that will not be rejected by one side or the other is to allocate the emissions' burden so that the date of convergence remains at seventy-five years.

The assumption that must be justified is that *maintaining the date of convergence is a focal point*. We do not have proof that this is so. We propose that it may be so. In fact, proposing this may help make it so. It is not only that the governments necessarily treat the date of convergence as of great political moment, although they might, but that the governments may believe that the date of convergence is highly salient to their polities, and a proposal that maintains that date is one that can be sold to the respective citizenries. More pessimistically, we can say that a proposal that entails a delay or an advance in the date of convergence will surely be vociferously objected to by one region's polity and so the *only* possibility for a solution is one that maintains the date of convergence.

The next step is to relax the assumption that the date of convergence is known by all. There is a fairly simple way of maintaining the date of convergence, even if we do not know what it is, as long as we can predict, fairly accurately, rates of growth in the near future, under various proposed schemes of burden sharing. The *growth factor* of an economy is one plus its growth rate. If growth rates are altered in two regions, but the *ratio of their growth factors* remains unchanged, then the ratio of their GDPs per capita will remain what it would have been without the alterations. For example, suppose that, under BAU, the Chinese growth rate would be 5% and the US growth rate would be 2% during the next year. The ratio of the growth factors is $1.05/1.02 = 1.0294$. This implies that the ratio of China's GDP per capita to the US's GDP per capita a year from now will be 1.0294 times what it is today. Now suppose China reduces its growth rate to 4% this year and the United States reduces its growth rate to 1.103%. One may compute that the ratio of growth factors, 1.04/1.0103, remains at 1.0294. This implies that the *ratio* of GDPs per capita will remain at what it would have been in the original scenario. If ratios of growth factors are invariant, then the date of convergence—which is the date at which the ratio of Chinese GDP per capita to US GDP per capita is unity—will remain unchanged.

We propose that the major regions of the world negotiate around the focal point of allocating emissions rights so as to maintain the ratios of their growth factors as invariants. If this proposal is followed, it will not reduce negotiations to a triviality—there will still be much to truck and barter about as regions argue about what their BAU growth factors would be and what the effects on growth factors would be of restricting GHG emissions. But the problem would be reduced to one whose solution international teams of economists could eventually agree upon. It would be reduced from a huge political problem to a rather technical one of economic computation.

This proposal motivates what we do in Chapter 5. We take the back-of-the-envelope estimate to be true that, under BAU, China and the United States would converge in seventy-five years—that is, in three generations of our model—and we ask: how should the global budget of carbon emissions, specified so that carbon concentration follows our path that converges to a level below 450 ppm, be allocated in each generation over the next seventy-five years, *in an optimal fashion*, so as to maintain that date of convergence of the *welfare levels per capita* in the two regions? We compute the solution to this problem.

We refer the reader to Chapter 5 for the details and state here what is perhaps the most important conclusion. We show that on the optimal sustainable path, the North's utility grows at about 1% per annum over the next seventy-five years, and the South's utility, after growing at 1% per annum during the first twenty-five years, grows at an average of 3.27% per annum in the fifty following years, catching up with the North in 2085.[10] Indeed, there are no feasible paths—that is, ones that obey all our constraints plus

[10] Some readers may be startled by the relatively low growth of the South's utility during the first twenty-five years. This result derives from the aggregation at the generational level and our choice largely to increase the utility of Generation 2 in the South at the expense of a small sacrifice in the utility of its Generation 1. Section A5.3 in Chapter 5 presents an alternative transition path.

the constraint of maintaining the date of convergence—at which the North can grow noticeably faster than 1% per annum on average.

There is both an optimistic and a pessimistic slant to this conclusion. The optimistic slant is that it is possible for the South to catch up to the North in seventy-five years *and* for the North to continue growing, while holding carbon concentration to an ultimate level below 450 ppm. Indeed, the South need not delay its catch-up date to the North because of the climate-change problem. This, we believe, is very good news. The global South need not give up on its goal of reaching the high standard of living of the global North. The pessimistic slant is that Northern citizens will have to accept a reduction of growth rates to about 1% per annum. And, similarly, the South (in particular, China) will have to accept a reduction in its growth rate that is considerable. This will be a major challenge for the politicians in both regions. US politicians will claim they cannot convince their citizens to settle for this slowdown, and the Chinese leadership apparently views its continuation in power as dependent upon continuing the high rates of growth that have characterized the last thirty years. Again, we believe the challenging issues are ones of politics, not of economic feasibility.

To summarize Chapter 5, we claim that the climate-change bargaining problem cannot be solved unless it is admitted that the issues of economic growth and the allocation of rights to emit are *solved together*. They are inextricably intertwined, and it is hopelessly naïve to suppose that one problem can be addressed in isolation from the other. Governments currently take the posture that their own economic decisions are not on the bargaining table, but if this posture is maintained, there is no solution to the challenge that faces us, and we must prepare to live on an increasingly hot planet, with all that that entails. We can only hope that presenting our analysis as clearly as possible will induce the relevant parties to cooperate with each other.

Chapter 6 explores, within simple theoretical models in the spirit of Chapter 2, two departures from the assumption, maintained until now and inherited from the *Stern Review*, that the probability of catastrophe at each generation, or hazard rate, (1) is exogenous, and (2) produces the extinction of the human species.

First (Section 6.2), we propose, more sensibly, that the probability of catastrophe is not appropriately conceived as that of a meteor colliding with Earth but rather as that of a climate-induced catastrophe, following from allowing an imprudently high concentration of atmospheric carbon to evolve. The assumption that the hazard rate is independent of our choices concerning GHG emissions is a poor one: adopting it is akin to looking for the lost jewel under the street lamp because that's where one can see the pavement.

The reason that researchers have assumed an exogenous hazard rate is that it immensely simplifies the calculation of optimal paths. In Chapter 6, for the first time in the book, we propose a *~~damage~~ function*, which relates atmospheric carbon concentration to the hazard rate at a point in time. We do not know how to estimate the ~~damage~~ function with any precision, nor do we think anybody does, at present. Nevertheless, it turns out that we are able to deduce some characteristics of the optimal path, which flow from characteristics of the ~~damage~~ function, a parametric form that we assume for the sake of argument. Formally speaking, Section 6.2 works with a very simple IAM, which contains the laws of motion of the economy; the relationship between economic variables and carbon concentration; and the relationship between carbon concentration and human welfare, as captured in the probability of a climate-induced catastrophe.

It is now important to observe that the pure-sustainabilitarian Ethical Observer places *no value* on the length of time that the human species exists. A human society that lasts for one date, where the single generation has a utility of 100, is indifferent, for this Ethical Observer, to a society that lasts for a thousand generations, where each generation's household enjoys a utility of 100. When the length of the species' lifetime was exogenous, this was of no import, because economic decisions had no impact upon the expected tenure of the species, but now that that tenure will be a result of the path chosen, the Ethical Observer's indifference to its length is of paramount importance.

Consequently, in Section 6.2, we adopt a new objective function, which could be called 'sustainabilitarian with a concern for species lifetime.' The von Neumann–Morgenstern (VNM) utility the Ethical Observer derives from a world that lasts for T generations is now assumed to be $T \min\{u_1, \ldots, u_T\}$. Thus, the Ethical Observer cares not only about the sustained *level* of utility but also about the sustained *length of time* of utility. The expected utility of this Ethical Observer is the sum of the VNM utilities she would have should the species end at any date along a proposed path, weighted by the probabilities that those events occur, where the probabilities themselves are endogenous to the choice of path.

We take a simple form for the damage function, which relates atmospheric carbon concentration to the probability of catastrophe; we assume that the probability of the species becoming extinct at any date is a weakly increasing function of the atmospheric carbon concentration at that date. This function is assumed to be constant (nonincreasing) until a certain level of carbon concentration is reached—perhaps the level of 350 ppm. But after this critical level is reached, the probability begins increasing at an increasing rate. We show, by simulation, that if the rate of increase of the damage function is sufficiently rapid, then the optimal path requires maintaining a carbon concentration very close to the critical level, which we take to be 350 ppm.

In particular, we compute the following: Suppose that we take the hazard rate per generation in the preindustrial era to have been 0.00017, or 0.017% per generation, as we proposed earlier. Now suppose that a doubling of concentration from 350 ppm to 700 ppm would double the hazard rate. Then our simulations show that the optimal path must maintain a concentration close to 350 ppm for a large interval of values of the key parameter in the damage function.

Now we must append to this conclusion the usual caveats: the model is simple, and we have made assumptions about the functional form of the damage function that may be wrong. Nevertheless, we believe that the simulations present an argument that should not be ignored.

The penultimate section, Section 6.3, introduces a second conceptualization of catastrophe. Instead of its being conceived of as the disappearance of the species, we now take catastrophe to be the occurrence of a tipping point, above which human life continues to exist but in a compromised state. We model this by assuming that the production function takes a beating at the tipping point: it becomes much more difficult to produce commodities, but human generations continue to exist at a diminished level of existence. Under this supposition, we calculate the optimal path of resource allocation. However, we assume in this section that the hazard rate is exogenous. Introducing one complication in the model is enough, at least for now. We model the sustainabilitarian Ethical Observer, in this case, as desiring to maximize the minimum *expected utility* over all generations, where each generation faces a lottery, *ex ante*, of living before or after the stochastic tipping point is reached. It turns out, perhaps unsurprisingly, that early generations should save to help generations that live after the tipping point occurs, and the greater the probability of that occurrence, the more the early generations should save. On the optimal path, every generation has the same expected utility, and so the sustainabilitarian ethic, once again, produces intergenerational equality. It would, of course, be possible to study, as well, paths that enable a constant growth rate in expected utility, analogous to our earlier growth-sustainability analysis.

The Conclusion summarizes our main conclusions.