CHAPTER

2 The economics of sustainable development

2.1 Introduction

'Sustainable Development' has become a political buzzword since the 1992 Rio Conference on the Environment, organised by the United Nations. But what exactly does it mean? That is a difficult question to answer since sustainable development (SD from now on) means different things to different people: people place varying emphases on multiple aspects of the rather vague notion which is SD. The best-known definition is that given by the Brundtland Commission in 1987 (WCED, 1987):

> development that meets the needs of present generations without compromising the ability of future generations to meet their own needs.

Another definition was offered by Asheim (1994):

> A requirement to our generation to manage the resource base such that the average quality of life we ensure ourselves can potentially be shared by all future generations

Many people see sustainable development as in fact serving multiple goals – economic development, a better environment and a particular concern for the poor (Pearce and Atkinson, 1998). At the more general level, development is indeed viewed rather differently to growth, implying some progress in areas such as health and education, rather than just increasing incomes. Two common features of many definitions of sustainable development are equity across and within generations, but most of the economics literature on sustainability has emphasised the former. Economists would say that SD is indeed principally an equity rather than an efficiency issue. However, the bigger the economic pie (broadly defined as total 'quality of life'), the more of it there is to go around, since other things being equal, economic growth raises the average level of well-being. There is thus a potential complementarity between promoting both effici and equity. This view is controversial: some would argue (e.g. Daly, 1990, Meadows *et al.*, 1992) that economic growth is in itself the *cause* of declining sustainability, since an expanding economic scale pushes increasingly against environmental constraints, threatening the operation of the joint economic–environment system, whilst there is a view that increasing income has a weak impact on increasing well-being above some threshold level.

It is also possible to distinguish between the idea of sustainability, namely the property of any system whose *performance* can be maintained over time, and sustainable development, namely the extent to which *development* can be sustained. In general, however, the terms 'sustainable development' and 'sustainability' are used interchangeably in economics, with each being about as contentious as the other. This chapter tries to pick out the main contributions that economists have made to the debate on sustainability, and on how we might measure it.

Why should we worry about the well-being of future generations? Two lines of reasoning have emerged. The first takes a utilitarian approach, and says that social welfare is concerned with the discounted sum of well-being of all people in a society over time. What the discount rate should be in calculating these present values is not something we have room to discuss here – the interested reader is referred to Sheraga and Sussman (1998) and Wietzman (1998). An alternative argument is that future generations have moral rights to a level of well-being, perhaps no less than our own. This Kantian view is in philosophical opposition to the utilitarian view above. Norton (2002) gives a good account of the issues involved in taking this position.

Economists' views on what *defines* a sustainable development path for an economy over time may be divided into two broad groups. The first (the outcome approach) is concerned with how the economic process directly affects human well-being. 'Well-being' is synonymous with the standard economic concept of utility or welfare of an individual. Hence, sustainability can be defined as the utility of a representative agent in any period t, $U(t)$ – taken to represent society's interests – to be non-declining for the rest of the time from time t^* onwards (Pezzey, 1992).

$$\partial U(t)/\partial t \geq 0, \qquad \text{all } t > t^* \tag{2.1}$$

or that in any period t, the utility of that representative agent does not exceed the maximum sustainable level of utility, depending on the economy's potential at time t (Pezzey and Toman, 2002):

$$U(t) \leq U_m(t) \tag{2.2}$$

where, for time periods s following on from time periods t:

$$U_m(t) = \max U \quad \text{given } U(s) \geq U(t) \qquad \text{for all } s \geq t \tag{2.3}$$

Equation (2.1) says that SD occurs when utility per capita is not falling over time (this means that constant utility equals sustainability, as well as rising utility). A variant on the outcome-based approach is to define SD in terms of the observable determinants of utility. In other words, if we know what factors affect utility – for example, the level of consumption and the level of environmental quality – then by examining changes in these factors we can infer whether a sustainable path is being followed according to Pezzey's definition. This idea is taken up again in Section 2.4, where we set out a formal model of sustainability indicators. A time path of consumption over time which is rising and then falling might be consistent with maximising the present value of social well-being, according to the utilitarian view, but would not be a sustainable path. There is thus a trade-off here between sustainability and present-value maximising optimality. Note also that the outcome definition of sustainable development implies that it is the absolute level of consumption and environmental quality per capita that matters for well-being, not one's consumption/environmental quality relative to one's neighbours, implying that rising real incomes result in higher utility. For a theoretical discussion of this assumption, see Pezzey (1997); for a recent review of empirical evidence, see Blanchflower and Oswald (2004).

The second economic approach to defining sustainability focuses on maintaining the means which are available to society to generate well-being or consumption, namely its resources. Resources consist of physical stocks and the technology which we use to exploit them. Economists have thought about SD from this viewpoint in terms of the concept of capital. Four forms of capital may be distinguished (Hanley and Atkinson, 2003):

1. *Man-made, or produced, capital*, K_m: This is the 'capital' that most economics students are familiar with. It comprises the results of past production, as the excess of output over consumption. K_m includes factories, machinery, roads, bridges, phone networks and satellites, and may be used up (depreciated) in the production of consumption goods and services. This depreciation needs to be offset with new investment or else the stock of K_m will decline over time.
2. *Human capital*, K_h: Human capital is people, their skills and knowledge. The stock of K_h can also depreciate (e.g. if unemployed people lose their skills), and can be added to through training and education.
3. *Natural capital*, K_n: Natural capital comprises all gifts of nature, and so includes renewable and non-renewable energy and material resources, clean air and water, nutrient and carbon cycles and biodiversity. Natural capital can clearly be depreciated when, for example, a non-renewable resource such as oil is used up or when a species dies out. Investments in K_n would include forest re-planting and re-stocking of fisheries.
4. *Social capital*, K_s: Recent attention has been directed towards the link between 'social capital' and sustainability, see World Bank (1997) for a discussion. Putnam (1993) speaks of social capital as comprising certain features of social organisation – norms of behaviour, networks of interactions between people and between institutions and trust between people. This could be important for sustainability in several ways. First, it is argued that there is an 'economic pay-off' from social capital whereby conditions favourable to economic growth are fostered by

a climate of trust between agents (Knack and Keefer, 1997). Second, there could also be an 'environmental pay-off' whereby, for example, strong community ties help enforce ownership regimes and management systems for common property (Grafton, 2000).

Sustainable development, in what we might refer to as an 'opportunities approach' to the issue, can then be defined as a non-declining stock of capital over time. Or, we can link the two approaches, and we show later that a declining stock of capital means development is unsustainable by definition (2.1), so that these two approaches overlap considerably, even if they are not identical.

2.2 Can development be sustained?

Economists have long been concerned with the question as to whether the natural environment imposes binding constraints on economic development and growth[1]. This concern starts with the writings of the classical economists Malthus, Mill, Ricardo and Jevons, who all recognised the possible limits that finite land supply, variable land quality and finite natural resources could impose on economic growth. Just over 100 years later, concerns re-emerged with the publication of *The Limits to Growth* (Meadows *et al.*, 1972), with worries over limits implied by environmental pollution being added to those of resource scarcity. The economists' response was to re-consider natural resource constraints, this time in the context of optimal growth models (Dasgupta and Heal, 1974; Solow, 1974). This work showed that, in the absence of technological progress, optimal growth patterns dictated an ultimately falling level of consumption and utility for future generations: in other words, optimal consumption would be single-peaked. A sufficient degree of (exogenously determined) technological progress coupled with 'enough' substitution possibilities between the different inputs used by the economy to produce consumption goods could offset this decline, but the trade-off between an optimal path and a sustainable one had been pointed out. Consumption can be sustained over time, but this implies a different objective function for society. None of this work, however, was set in the explicit context of sustainable development. More recently, the development of ideas of endogenous technological progress has led to a re-evaluation of the links between resource depletion, population growth and economic growth, and has on the whole led to more optimistic conclusions on the degree to which growth can be sustained (Bretschger, 2005), particularly where the sector of the economy that 'produces' new ideas is less resource-intensive than other sectors.

An important move towards an economics of SD was to show conditions under which natural resource scarcity was consistent with non-declining consumption for an

[1] For an excellent overview of the history of economists' concerns with sustainable development, see Pezzey and Toman (2002). This section draws heavily on this source.

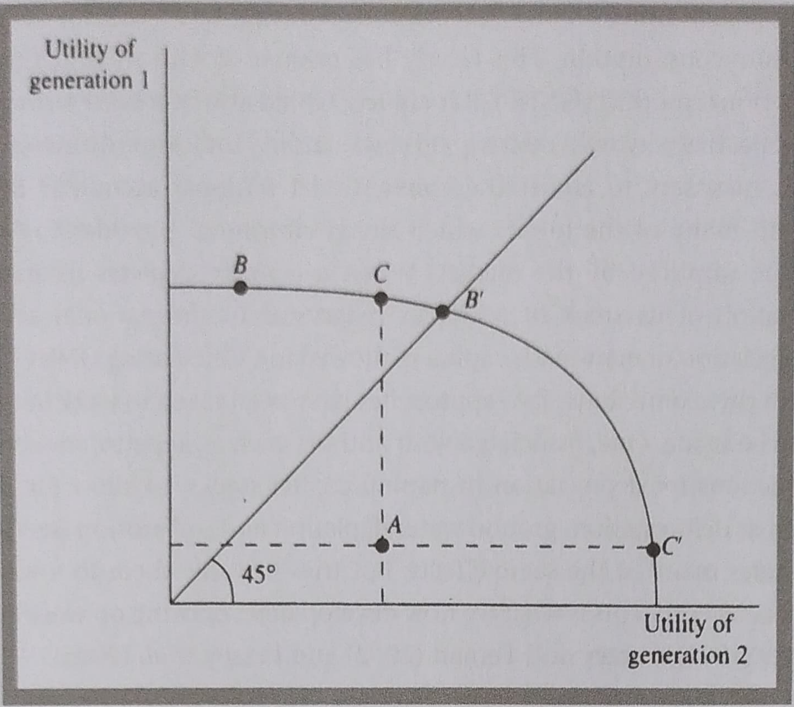

Figure 2.1 Sustainability versus efficiency

utility (welfare) of generation 2 has been reduced below that of generation 1. This will be so unless costless transfer mechanisms enable us to redistribute utility from B to, say, B′. Points along the line segment B′C can only be chosen if we know what the intertemporal social welfare function looks like.

2.4 Economic indicators of sustainability

In the following section, we set out the two main economic indicators of sustainability, namely green net national product and genuine savings. Both are firmly based on the weak sustainability assumption. Only one 'version' of these indicators will be presented here, and is due to Pezzey and Toman (2002) and Pezzey *et al.* (2006). Summaries of the very considerable literature on these indicators can be found in Hanley (2000), Pezzey and Toman (2002) and Asheim (2003). The section concludes with a brief look at strong sustainability indicators.

2.4.1 Green Net National Product

A large literature has recently emerged on whether the well-known macroeconomic measure gross national product (GNP) can be transformed to produce an indicator of SD (Asheim, 2003). GNP has traditionally been thought of both as a welfare measure and as a measure of national income. By relating this to a particular concept of income put forward by Hicks in 1946, some authors have sought to transform GNP into an indicator of SD. Hicks' view on income was that it represented that portion of the value of output which could be consumed in any year without reducing one's wealth, defined as one's potential future consumption. This clearly has resonance with some definitions of SD. Green net national product (GNNP), it is argued, would also be a *better welfare measure* for society – irrespective of whether it is a good sustainability indicator (Aronsson *et al.*, 1997).

Why is it necessary to adjust the conventional national accounts? Because these accounts omit many of the inputs which the environment provides to the economy, since they are unpriced by the market. When a country depletes its natural capital, this depreciation of its stock of resources is ignored in the national accounts, even though depreciation of man-made capital is allowed for. Calculating GNNP thus involves correcting for these omissions. Two approaches have been taken to work out what corrections should be made. One, associated with authors such as Repetto, involves a series of ad hoc deductions for depreciation in natural capital stocks, to allow for development impacts such as deforestation, ground water depletion and soil erosion (see Box 2.1). The second includes many of the same effects, but tries to value them in a way consistent with economic theory. This is what we now develop here, drawing on work by Weitzman (1976), Maler (1991), Pezzey and Toman (2002) and Pezzey *et al.* (2006).

BOX 2.1

Natural resource accounting in Asia

A country can fell its forests, erode its soils, exhaust its minerals, pollute its aquifers and erase its wildlife, all without adversely affecting its measured income. By failing to recognise the asset value of natural resources, the UN System of National Accounts (SNA) misrepresents the policy options which nations face. In Costa Rica increases in the rates of deforestation, soil erosion and the consequent impacts on inshore fisheries and coral reefs, compounded by overexploitation, have had major socioeconomic impacts. Natural resource accounts compiled for the 1970s and 1980s using remotely sensed data on land use change, field data on forest productivity, GIS (geographic information system) studies of soil erosion and sample studies of fish populations reveal the problem. These studies suggested that Costa Rica had been depleting its forest, soil and fishery capital by at least 5 per cent on average of GDP per year since 1970. The asset value and sustainable profits of the principal fish species in the major fishing area in the Gulf of Nicoya dropped to zero as fishermen's earnings fell below the level of welfare payments to the destitute. Leaving aside the unquantified service value of Costa Rica's forests (as wildlife habitat, tourist attractions, ecosystem regulators and suppliers of non-timber products), the forestry sector generated substantially negative net national income throughout the 1980s and overall deforestation estimates from 1966 to 1989 were 28.2 per cent. Ignoring declines in soil fertility due to losses of micronutrients, biological activity and desirable soil structure, and restricting off-site concerns to siltation effects on hydroelectric systems, economic losses, primarily through the cost of replacing lost macronutrients, accounted for 9 per cent of all agricultural production. Using these figures, it can be shown that natural resource depreciation rose from 26 per cent of gross capital formation in 1970 to 39 per cent in 1989. Thus the conventional accounting framework overstated actual net capital formation by 70 per cent in 1989 (Repetto, 1992).

Using oil and forestry data for 1970–84 from Indonesia and a Javanese soil erosion study in 1985, Repetto *et al.* (1989) presented national accounts which incorporate measures of natural resource depletion, yielding environmentally adjusted 'net' domestic product (NDP). Indonesia's national income and economic growth have been overstated by conventional GDP, which increased at an annual rate of 7.1 per cent from 1971 to 1984. Revenues from oil and gas production, hard mineral extraction and the harvesting of timber and other forest products were used to finance government development and routine expenditures. However, the losses of oil reserves, topsoil and forest cover have resulted in a depleted natural resource base which will restrict future development opportunities. The annual rate of increase of NDP was in fact only 4.0 per cent. The raw data used for the soil erosion study are being debated and recent data give reason to doubt the assumed magnitude of soil erosion and the importance of human influences on the process. Nevertheless, a more complete accounting system which estimates the depreciation in future productive potential for other environmental resources such as fisheries, non-timber forest products, natural gas, coal, copper, tin and nickel might reveal more unsustainable trends.

Consider an economy with a representative agent who derives utility from consumption of both produced goods and environmental amenities, given by a vector $\mathbf{C}_t$, where t indexes time. Production is determined by the aggregate (man-made plus natural plus human) capital stock, a vector $\mathbf{K}$, and technological progress which depends solely on the passage of time. An economy is deemed to be sustainable at time t if utility is less than or equal to maximum sustainable utility at this time, where 'sustainable' here means consistent with non-declining values of U(C), as in (2.4) and (2.5) . The economy maximises the present value of utility over infinite time, at a constant discount rate ρ:

$$\max_{C,K} \int_0^\infty U[C_t]e^{-\rho t}dt \tag{2.4}$$

Pezzey and Toman show that for this economy to be sustainable, green net national product Y^t, defined by

$$Y^t = P(t)\cdot C(t) + V(t)\dot{K}(t) \tag{2.5a}$$

where $\dot{K}$ is the rate of change in K per unit of time, subject to production possibilities given by K(t) and t and where P is the relative price for the consumption goods and environmental amenities and V is the price for each element of the capital stock, *must be non-declining at time t*, that is

$$\dot{Y}(t) \leq 0 \Rightarrow U(t) > U_m(t) \tag{2.5b}$$

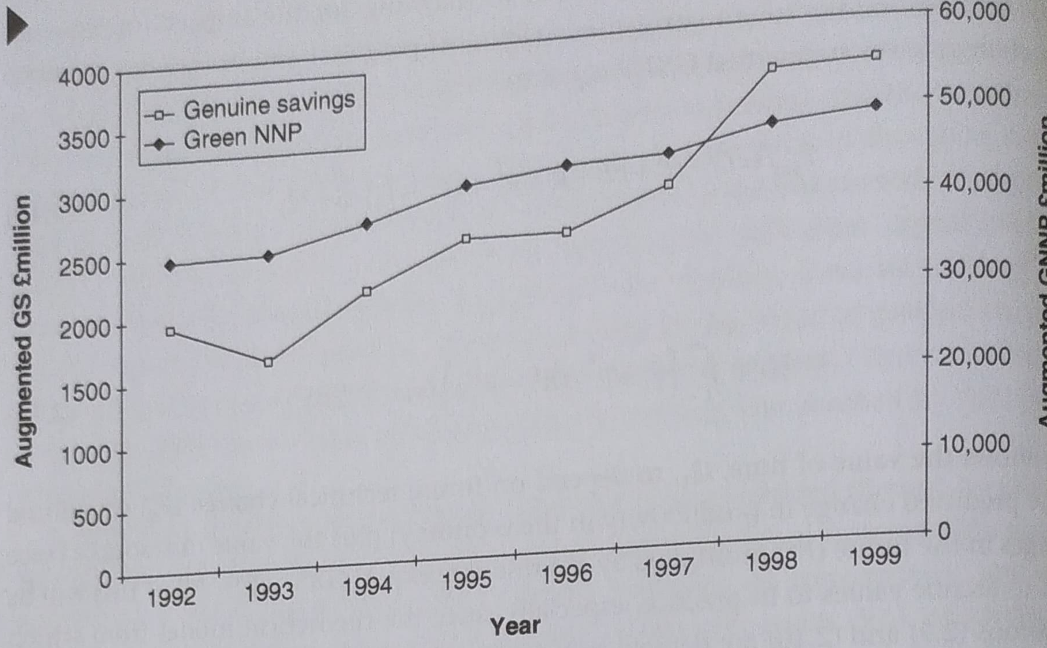

Figure B2.1 Augmented green NNP and genuine savings for Scotland

How good a measure of sustainability is the change in GNNP? Several weaknesses may be noted. First, there is disagreement among economists of how exactly adjustments to the conventional accounts should be made (e.g. over how to treat new discoveries; and how to value environmental liabilities, see Hanley (2000)). It is also true that, for the adjustments to be correct, the price and marginal cost values used should be those which result from a competitive, dynamically optimal use of resources. However, the well-known problem of the absence of clearly defined property rights means this is unlikely to be true, especially for fisheries. There is also a dispute amongst economists about whether GNNP can be used even in principle as a sustainability indicator. Dasgupta *et al.* (1996) suggest in a less restrictive model that if a correctly evaluated GNNP is rising then the country's long-term welfare is increasing. However, Asheim (1994) shows that unless the adjustments to conventional GNP are made using the prices that would hold along a *sustainable* (as distinct from an inter-temporally efficient) path, GNNP tells us nothing about sustainability. Consistent with this, Pezzey and Toman (2002) showed that GNNP is only a one-sided indicator of SD: it only tells us when an economy is behaving *un*sustainably, but cannot guarantee that an economy is in fact sustainable.

2.4.2 Genuine savings

An alternative economic indicator of SD, closely related to GNNP, is the *genuine savings* concept, put forward originally by Pearce and Atkinson (1993) and Hamilton (1996). GS compares investment in an economy with depreciation of all forms of capital. As such, it is an empirical application of the 'opportunities' approach to sustainability, and has been thought of as a test of whether an economy is following the Hartwick rule. Since it involves a comparison of investments and depreciation in total capital, it is also referred to as *net investment*. How is GS calculated? Again, contrasts can be drawn between sometimes ad hoc empirical approaches and approaches based on underlying theory. In fact, GS can be derived from exactly the same theoretical model as that outlined above for GNNP. The equivalent GS to GNNP shown in (2.9) is

$$GS = V \cdot \dot{K} = \dot{K} + \dot{K}^{f} + (P^{R} - f_{R}) \cdot \dot{S} \tag{2.12}$$

that is, GS is given by the net change in produced capital K, plus the change in foreign capital, plus the value of depreciation of natural resource stocks, given by the physical change times the Hotelling rent. A time-augmented version of GS can also be given, which would be just (2.12) modified for a value of time term Q^t similar to that outlined for GNNP.

The test of unsustainable development is then whether GS is negative or not. That is,

$$GS(t) < 0 \Rightarrow U(t) > U_m(t) \tag{2.13}$$

We can include changes in the stock of human capital in Equation (2.12) – spending on education and training would be a positive term here – whilst augmentation of the indicator with the value of time Q^t allows for technical change to be included. If GS is negative, then this is a clear indication of unsustainable behaviour. Investments in the aggregate capital stock must be increased. This increased investment could include investments in pollution treatment, fisheries and forest stocks, and searching for new, economically viable, non-renewable resource deposits. It could equally comprise increased investments in hospitals, schools and railways. The GS indicator has recently been extended to allow for international trade effects (Proops *et al.*, 1999, see Box 2.5). GS and GNNP are closely linked to each other since they are both derived from the same underlying theory. It is possible to show formally that (2.5b) and (2.13) are equivalent tests of sustainability *in theory*: if GS is negative, then GNNP must be falling. Furthermore, Asheim and Weitzman (2001) show that the rate of change of GNNP at time t must be equal to the real interest rate multiplied by GS. Pezzey and Toman (2002) show that this relationship holds for the time-augmented versions of the two measures. If this is so, then which measure we prefer may come down to which data it is easier to acquire: the data needed to measure GS or the data needed to measure GNNP. Since the former is less than the latter, this suggests a preference for GS.

Box 2.3 shows some results from Hamilton and Atkinson's (1996) estimates of GS for the United Kingdom for the period 1980–1990. Results from the Hanley *et al.* (1999) study for GS for Scotland are also given in Box 2.3. The Scottish economy can be seen to have been behaving unsustainably for much of this period, since GS were negative. This mirrors the finding by Hamilton and Atkinson (1996) for the UK as a whole over this period. However, the discrepancy between depreciation of capital stocks and savings starts to decrease towards the end of the period.