6 Energy Reserves and Sustainability

The finiteness of fossil energy sources gives rise to the question of whether sustainable economic development is possible at all since these resources will increasingly become scarce and even cease to be available. Resource economics—the theory of dealing with the efficient use of exhaustible resources—has been addressing this problem. Grounded in the pertinent economic models, this chapter revolves around the following issues:

– How are energy reserves measured and how large are they?
– What is the optimal extraction strategy for the owners of an exhaustible resource?
– What is the optimal rate of extraction from a welfare point of view?
– Does market failure occur, i.e. are there systematic deviations from the optimal extraction path?
– What are the consequences of the increasing physical scarcity of energy sources for the price of energy?
– How far can these prices rise?

However, the optimum from the point of resource economics, while resulting in an efficient use of exhaustible energy resources, need not be sustainable. The relationship between economic efficiency and (so-called weak) sustainability therefore needs to be clarified. This leads to additional questions:

– What are the conditions that make sustainable development possible in spite of continued use of non-renewable energy sources?
– For instance, does the global oil market satisfy these conditions?
– What interventions might be called for in order to satisfy the conditions for weak sustainability?

The variables used in this chapter are:

P. Zweifel et al., *Energy Economics*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-662-53022-1_6

C	Consumption
c	Marginal unit cost
$Disc$	Discovery of energy reserves
δ	Depreciation rate
f	Rate of autonomous technological change
H	Hamiltonian function
η	Income elasticity of energy demand
i	Market interest rate
K	Reproducible capital (as distinguished from natural capital)
L	Lagrange function
λ	Opportunity cost of resource extraction (also called scarcity rent), a Lagrange multiplier
NPV	Net present value of cash flows
Π	Profit
Q	Production function
p	Price
R	Extraction of energy reserves
r	Social time preference
S	Stock of reserves
σ	Elasticity of substitution
T	Planning horizon
U	Utility function
W	Welfare (wealth), value of the objective function
ω	Opportunity cost of consumption, a Lagrange multiplier

6.1 Resources and Reserves

Whether a raw material constitutes a resource or not depends on available technological know-how and the capability of using it. For example, the uranium isotope U_{235} became an energy resource only after control over nuclear fission was achieved. Next, 'resources' have to be distinguished from 'reserves' (see Table 6.1).

Resources comprise all useful raw materials existing in the ground, including those whose deposits are only presumed to exist or are currently too costly to be extracted using available technologies. Reserves are those resources that exist with high probability and can be extracted at a cost below their market price. Accordingly, higher market prices can cause resources to become reserves. The same holds for increased efforts at exploring an assumed resource deposit and lowering the cost of a mining or extraction technology. Adding cumulated amounts extracted to the total stock of reserves and resources leads to an estimate of ultimately recoverable resources.

Table 6.1 Ultimately recoverable resources

Cumulated extraction	Remaining reserves		Resources that are not (yet) reserves	
Cumulated extraction	Physically proved, technically feasible and economically viable		Proved but technically not feasible and/or economically not viable	Not proved but possible according to geological evidence
	Certain	Probable		
		←P→		

Source: Erdmann and Zweifel (2008, p. 122)

Following the World Petroleum Council, reserves are classified according to the probability of economically viable extraction (see Campbell and Laherrere 1996):

– P (proved): probability of extraction >90%;
– 2P (proved + probable): probability of extraction >50%;
– 3P (proved + probable + possible): probability of extraction >10%.

The U.S. Securities and Exchange Commission as well as other financial regulators require that resource-extracting companies listed on the stock exchange use P-reserves when reporting their assets, whereas geologists as well as internal planners of resource companies use 2P-reserves as the relevant figure. As a consequence, published reserves can be higher merely due to reclassification of known deposits rather than new discoveries (resulting in so-called paper barrels). Indeed, experts such as the Texan investment banker Simmons claim that increases of oil reserves in recent years have been more due to such reclassifications than to the discovery of new deposits.

6.1.1 Resources

There is a multitude of publications concerning globally available energy resources. Estimates sometimes diverge substantially, as they are based on data provided by resource-extracting companies, governments, and independent experts.[1] According to Table 6.2, the oil resources amount to 470 bn toe or 3400 bn bbl, respectively.

Several models have been developed for determining the amount of an oil and gas resource that can be ultimately recovered. One of them is by geologist Hubbert (1956, 1962), stating that accumulated discoveries follow a logistic trajectory (see Fig. 6.1). Thus, when the industry is in its start-up phase, only few oil deposits are discovered. With more experience (and therefore decreasing marginal cost of exploration) the rate of discovery increases. However, when the bulk of existing

[1]Some experts doubt the credibility of data published by several OPEC countries, who have been stating constant reserves of crude oil for many years.

Table 6.2 Global fossil energy reserves and resources 2013

Energy source (bn toe)	Cumulated extraction up to 2013[a]	Reserves 2013[b]	Resources which are not yet reserves, 2013
Conventional oil	175	244	231
Unconventional oil[c]		69	392
Conventional gas	85	250	414
Unconventional gas		7	686
Coal	134	697	16,747
Uranium	18	21	228
Thorium			109

[a]BP (2014)
[b]BGR (2014)
[c]Hydrocarbons not capable of flowing

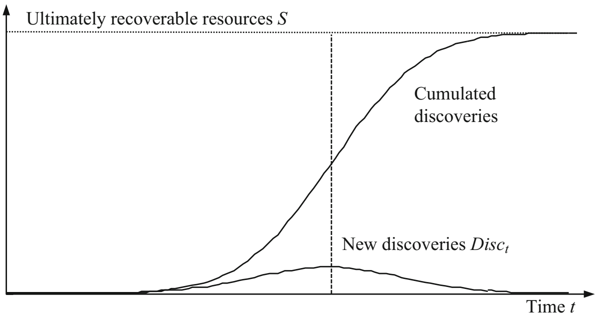

Fig. 6.1 Logistic path of cumulative global resource discoveries

deposits has been located, the rate of discovery falls again. Accordingly, cumulated discoveries increase more slowly, approaching the limit of ultimately recoverable resources.

If one assumes symmetry between the phases of increasing and decreasing rates of discovery $Disc_t$, cumulated discoveries up to time T follow the logistic equation

$$\int_0^T Disc_t\,dt = S\frac{1}{1+\exp[-a(T-t_{max})]} \quad \text{with}\, a > 0 \tag{6.1}$$

Here, t_{max} denotes the year where the rate of discovery reaches its maximum and S, total reserves. It can be shown that t_{max} determines the time when one-half of the reserves have been discovered. Differentiation with respect to time yields

$$Disc_t = aS\frac{\exp[-a(t-t_{max})]}{(1+\exp[-a(t-t_{max})])^2} \tag{6.2}$$

Econometric methods can be applied to derive Maximum Likelihood estimates of the unknown parameters a, t_{max}, and S from time series data on $Disc_t$ and t.

Therefore, the history of discovery rates permits to infer the unknown resource stock *S*—provided Eq. (6.1) holds.

From an economic point of view, however, this model is weak because it neglects exploration costs, technological change in exploration, resource prices as well as institutional conditions such as public versus private ownership of reserve deposits (Kaufmann and Cleveland 2001; Reynolds 2002). Moreover, the assumed symmetry between increasing and decreasing rates of discovery does not conform to reality.

6.1.2 Static Range of Fossil Energy Reserves

In spite of these qualifications, estimates of ultimate recoverable oil resources published in the literature seem to converge (see Fig. 6.2). This convergence may be the result of two opposing economic forces. On the one hand, reserves amount to an intermediate product since investment must be made in the exploration of deposits, in the purchasing of extraction rights, and in the enforcement of already acquired property rights. With decreasing global reserves, the price of oil is expected to increase, making it economically interesting to invest into the creation of additional reserves. Thus, changes in global reserves crucially depend on the expected value of exploration costs compared to expected oil prices. On the other hand, accumulated exploration leads to learning effects (see IEA 2000; Nakicenovic 2002), which are a major source of cost reductions. High oil prices and stepped-up exploration efforts have resulted in innovations (such as 3D seism in the 1990s and 4D seism and fracking since 2000) that serve to lower the cost of exploration. As stated by Adelman (1990), there is a permanent race between the decrease in the reserves remaining and the increase in technological knowledge.

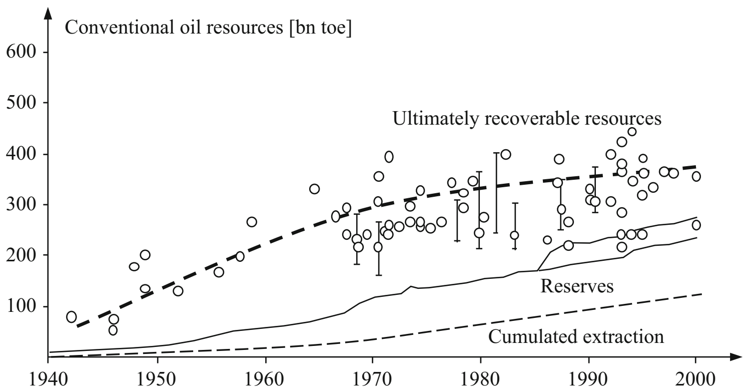

Fig. 6.2 Discovery of conventional oil resources over time. Source: Erdmann and Zweifel (2008, p. 125)

Reserves constitute an asset that is necessary for long-run production planning, securing deliveries, and market presence. Moreover, they help a resource-extracting company to attain a favorable credit rating. However, the returns on the investment needed for the acquisition of this asset have low net present values if extraction of a new resource deposit starts far in the future (see Sect. 3.1). From an economic point of view, exploration efforts should be exerted to the point where their marginal cost is still covered by the increase in the present value of expected returns on the oil found. As a consequence, an individual resource-extracting company is predicted to have reserves that are neither excessive nor insufficient. Whether reserves are excessive or insufficient depends, among other things, on the future expected price of oil. However, the price of oil not only determines the individually optimal amount but also the global amount of resources. Higher (expected) oil prices tend to increase the global amount of resources, at least in the long term.

A common indicator of remaining reserves is time to depletion (also called range). In its static variant, this is the ratio of remaining reserves over the current rate of extraction. In its dynamic version, the change over time in both quantities is accounted for. Figure 6.3 shows the static range for conventional crude oil and natural gas reserves. It is roughly constant during the past decades, in spite of increasing extraction rates of oil and gas. This constancy corresponds quite well to the optimality of investment in exploration discussed above. Moreover, the static range of natural gas consistently exceeds that of crude oil. This can be attributed to the fact that large gas discoveries often occur as a byproduct of oil field exploration (so-called associated gas).

However, the static range does not inform about how long the energy source will still be available. For one, reserves increase due to exploration and reclassification of reserves. The effective time to depletion can exceed the static range if the rate of annual extraction falls—and *vice versa*. Nonetheless, the static range is a helpful indicator of global reserves available. A decline in its value should be seen as indicating a need for increasing investment in exploration and efforts to substitute the resource by other (renewable) energy sources.

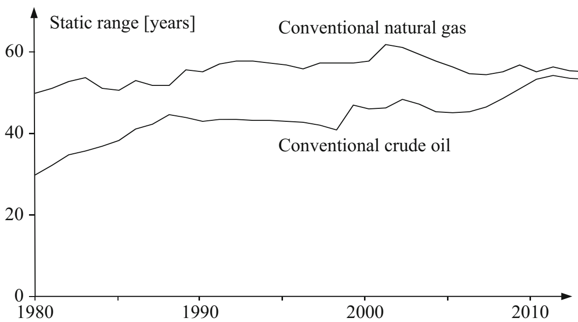

Fig. 6.3 Static range of conventional oil and natural gas reserves. Data source: BP (2014)

6.2 Profit-Maximizing Resource Extraction

In the same way that the transformation from resources to reserves is governed by economic considerations, the transformation of a reserve into money (i.e. extraction and sale) is an economic decision. It revolves around two alternatives:

- Leave the reserve in the ground and wait for a higher market price (which is to be expected due to increased scarcity);
- Extract the resource and invest the profit in securities or assets thereby earning the market interest rate.

6.2.1 Hotelling Price Trajectory

The Swedish economist Hotelling (1931) found a solution to the decision problem stated above, i.e. the profit-maximizing quantity of an exhaustible resource to be extracted during a given period. His model introduces several simplifying assumptions, notably competitive and efficient markets for resources, reserves, energy sources, and capital, profit-maximizing resource owners, perfect information regarding the amount of reserves, constant cost of mining and extraction, no inflation, no inventories after extraction, and an intertemporally stable demand function depending on the price of energy only. Given these assumptions, profit Π_t in period t is given by

$$\Pi_t = p_t R_t - cR_t. \tag{6.3}$$

Here, $R_t \geq 0$ is the quantity extracted and sold in period t, p_t the market price, and c the marginal cost of extraction, which is assumed to be constant and hence equal to unit cost for simplicity.

Postponing extraction is advantageous as long as (expected) profit in the following period Π_{t+1} exceeds profit achievable in the current period t, invested at the (real) rate of interest i,

$$\Pi_{t+1} = p_{t+1}R_{t+1} - cR_{t+1} > \Pi_t(1+i). \tag{6.4}$$

In the opposite case, the firm extracts during the current period. If all extracting companies decide in this way, there will be an equilibrium market price p_t, determined by their decisions in the aggregate, which has to satisfy the equality condition,

$$\Pi_{t+1} = p_{t+1}R_{t+1} - cR_{t+1} = \Pi_t(1+i). \tag{6.5}$$

Iterating this idea until the end of the planning horizon T—and assuming perfect foresight with respect to price—leads to the following decision problem of reserve-owners seeking to maximize the net present value NPV of their asset,

$$NPV = \sum_{t=0}^{T} \Pi_t \cdot (1+i)^{-t} = \sum_{t=0}^{T} (p_t R_t - cR_t)(1+i)^{-t} \rightarrow \max! \tag{6.6}$$

The decision variables of each company are the extraction rates R_t during the planning period $t = 1, \ldots, T$. However, their optimization is subject to the constraint that the sum of extractions must not exceed total stock in the ground S. On the other hand, it would not make sense for the reserve-owner to leave anything in the ground beyond the planning period T so that the constraint becomes

$$\sum_{t=0}^{T} R_t = S. \tag{6.7}$$

This constraint can be integrated into the objective function using a Lagrange multiplier $\lambda > 0$,

$$L = \sum_{t=0}^{T} (p_t R_t - cR_t)(1+i)^{-t} - \lambda \left(\sum_{t=0}^{T} R_t - S \right) \rightarrow \max! \tag{6.8}$$

This expression states that if accumulated extractions were to exceed the existing stock S, the value of the objective function L would be reduced because of $\lambda > 0$, causing the degree of goal attainment to fall. Therefore, one of the first-order optimality condition reads

$$\frac{\partial L}{\partial \lambda} = \sum_{t=0}^{T} R_t - S = 0. \tag{6.9}$$

It guarantees that the constraint (6.7) is always satisfied in the optimum. The second optimality condition concerns the rate of extraction R_t during period t. Since the other extraction rates $R_0, \ldots, R_{t\text{-}1}, R_{t+1}, \ldots; R_T$ are not affected by a decision in period t, one obtains

$$\frac{\partial L}{\partial R_t} = (p_t - c)(1+i)^{-t} - \lambda = 0. \tag{6.10}$$

Solving this optimality condition yields the so-called Hotelling price trajectory,

$$(p_t - c) = \lambda(1+i)^t \,\text{or}\, p_t = c + \lambda(1+i)^t. \tag{6.11}$$

Equation (6.11) can be interpreted as follows.

– Marginal extraction cost c as lower limit on the price p: If the reserve were available in unlimited quantity, the constraint (6.7) would not be binding. From Eq. (6.8), one would obtain

$$\frac{\partial L}{\partial S} = \lambda. \tag{6.12}$$

Therefore, the Lagrange multiplier reflects how much the objective function—the present value of profit in the present case—would increase if the constraint were to be reduced by one unit (or if the reserve initially were larger by one unit, respectively). However, if the reserve is unlimited, a further increase does not contribute to the value of the objective function, implying $\lambda = 0$. In this case, the market price of the reserve is equal to the marginal cost of extraction, in keeping with Eq. (6.11). Thus, one obtains the classical rule, "price equal marginal cost" that characterizes competitive markets without exhaustible resources.

– Surcharge on marginal cost, scarcity rent, user cost: The surcharge on the marginal cost of extraction depends on the value of the Lagrange multiplier λ. From the above interpretation of Eq. (6.12), one may infer that λ has a large value when the stock S is small and *vice versa*. Thus, the surcharge is high and increases fast when the reserve is small, causing λ to be large. Note that only the initial value of S is relevant for depletion along the Hotelling trajectory.
– Price increase over time: Even given a constant marginal cost of extraction, the market price of the reserve increases over time. This can be seen by writing Eq. (6.11) for period $t-1$,

$$p_{t-1} - c = \lambda(1+i)^{t-1}. \tag{6.13}$$

Dividing (6.11) by (6.13), one obtains

$$\frac{p_t - c}{p_{t-1} - c} = 1 + i \text{ or, after multiplying by } (p_{t-1} - c), \quad \frac{p_t - p_{t-1}}{p_{t-1}} = i. \tag{6.14}$$

Therefore, the (inflation-adjusted, real) price of the resource increases exponentially at a rate that equals the rate of interest on capital markets. This increase reflects the fact that the reserve becomes increasingly scarce as extraction continues, requiring its market price p_t to go up relative to the prices of other goods and services.

The difference $p_t - c$ is called scarcity rent (or user cost, respectively). It amounts to the economic value of the reserve in the ground. If a company were to acquire the property right of a deposit, it would have to pay a price according to this user cost, provided markets for property rights are efficient.[2] Indeed, the sum of so-called nonreproducible capital (i.e. the reserve in the ground) represents total fixed assets

[2] In many countries exhaustible resources are considered public property, which hampers trade in deposits and causes pertinent markets to deviate from economic efficiency.

of a resource-extracting company. In the optimum, both types of assets (reproducible, nonreproducible) achieve the same rate of return.

6.2.2 Role of Backstop Technologies

No price can increase without limit. Sooner or later, the price of a reserve will attain p_{subst}, at which some energy source becomes competitive as a substitute. Assuming that this alternative source can supply an amount sufficient to match demand at that price, the sales price of the reserve-extracting industry cannot exceed p_{subst} because its product would not be competitive anymore. This alternative energy source is the so-called backstop technology that will substitute the reserve, at the latest once it is exhausted (and possibly sooner).

This fact determines the optimum supply price in the current period, which can be shown as follows. If extracting companies are successful in maximizing net present value, the scarcity rent in the last period ($p_T - c$) must be equal to ($p_{subst} - c_{subst}$). Discounting back to period t, one obtains[3]

$$p_t - c = (p_T - c)(1 + i)^{t-T} = (p_{subst} - c)(1 + i)^{t-T}. \tag{6.15}$$

More generally, the Hotelling price trajectory describes the stepwise transition from the least costly but scarce energy source to the next-best but more expensive substitute. This goes on until the most costly but unlimited backstop technology becomes competitive. In this process, the cheapest deposit is used up first, followed by a cascade of more costly deposits.

One last property of the Hotelling trajectory is noteworthy. Figure 6.4 shows how—depending on the rate of interest—optimal paths of the reserve extraction and the associated prices may look like given the assumptions made above. In this example, the initial stock of reserves is $S = 100$ and the price of the backstop technology, $p_{subst} = 80$. In all trajectories, the price of the backstop is reached, but the point in time depends on the rate of interest. At low interest rates, the optimal extraction period is long, the initial price of the reserve high, and the price increase over time relatively slow.

The reason is that at a low rate of interest, the opportunity cost of foregoing interest income by leaving the resource in the ground rather than selling it is low, causing the optimal extraction period to be long. Conversely, at a high rate of interest, this opportunity cost is substantial, calling for timely extraction. At extremely high rates of interest, it is optimal to deplete the deposit as quickly as possible. In that case, the owner of the extraction right does not take scarcity rent

[3]This is the typical approach for solving dynamic optimization problems: Calculation of the current optimal price is based on the optimal intertemporal price trajectory defined by the so-called transversality condition (6.14) which is a dynamic equation with at first unknown parameters. Based on the conditions for the final state, the optimal price given the parameter values can be determined, resulting from backward induction from future T to present t.

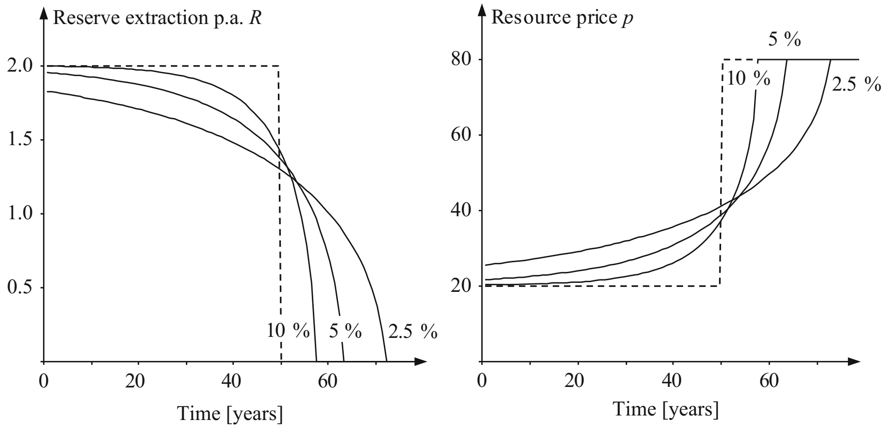

Fig. 6.4 Optimal extraction trajectories of an exhaustible resource

into account. This causes the remaining period of utilization to fall—by as much as two decades in the example of Fig. 6.4.

Another determinant of the optimal extraction period is the size of the cost advantage $(p_{subst} - c_{subst})$, also called differential rent. To illustrate, let T^* be the period in which $(p_{subst} - c_{subst}) = (p_T - c)$ exceeds 5% of marginal extraction cost c. In view of Eq. (6.15), T^* must satisfy the following condition,

$$(p_T - c)(1+i)^{-T^*} = 0.05 \cdot c \text{ or}$$
$$T^* = \frac{1}{\ln(1+i)} \cdot (\ln 20 - \ln c + \ln(p_T - c)). \tag{6.16}$$

According to Eq. (6.16), T^* does not depend on the stock of reserves S but on marginal extraction cost c, the cost advantage of the deposit, and the real interest rate i. The differential rent $(p_T - c)$ is determined, among other things, by the (expected) backup cost c_{subst} [see Eq. (6.16) again]. If research and development leads to a drop in this cost, the differential rent decreases, causing the present rate of extraction R_t to increase and the price of the reserve to fall. Thus, lower backup costs are beneficial to the users of the resource and society in general.

However, the investment in backstop technology pays back only after T^*, when it becomes competitive. Private investors will hardly undertake investments into research and development if T^* is far off. Yet such investments generate a positive externality, which may provide a justification for their public (co-)financing (IEA 2000).

6.2.3 Role of Expectations and Expectation Errors

One of the assumptions underlying the Hotelling price trajectory [defined by Eqs. (6.14) and (6.15)] is perfect foresight. However, parameters such as the size of deposits, their quality, the future cost of extraction, and future market prices are highly uncertain. One may use average expectations held by market participants to render the Hotelling trajectory applicable to the case of uncertainty (provided market participants are risk-neutral).

However, expectations may turn out to be wrong. If a majority of market participants have to adjust their expectations, reserve quantities supplied and prices must adjust as well. Table 6.3 illustrates the case of two crude oil deposits with differing marginal cost of extraction. If the owner of the extraction right expects a price increase of 2 USD/bbl, the scarcity rent of the deposit increases by 6.7% for deposit A and 10% for deposit B, respectively. Assuming an interest rate of 8%, deposit A would be depleted, since its scarcity rent increases more slowly than the interest rate ($0.067 < 0.08$). Deposit B would be preserved for later use ($0.10 > 0.08$). This corresponds to the general rule that low-cost deposits are exploited first.

Now let the expected price increase be 5 USD/bbl, causing the owner of the extraction right to expect an increase of the scarcity rent of 16.6% and 25%, respectively. In this case, it becomes optimal to defer extraction for both deposits. As a consequence, current market supply is reduced and price tends to increase.

This example illustrates the phenomenon of self-fulfilling expectations. When the owners of extraction rights expect the price of their reserve to increase, they reduce the rate of extraction, thus reinforcing expectations. Since price expectations may change fast, for example due to revised estimates of reserves, resource markets are predicted to exhibit excessive price volatility. Indeed, the standard deviation of daily crude oil prices is high, amounting to some 30% according to Plourde and

Table 6.3 The role of expectations: a crude oil example

	Deposit A	Deposit B
Price in period *t* (USD/bbl)	50	50
Cost per barrel (USD)	20	30
	Expected price increase 2 USD/bbl	
Scarcity rent (USD/bbl)		
In period *t*	30	20
In period *t*+1	32	22
Growth rate of scarcity rent (%)	6.7	10.0
	Expected price increase 5 USD/bbl	
Scarcity rent (USD/bbl)		
In period *t*	30	20
In period *t*+1	35	25
Growth rate of scarcity rent (%)	16.6	25.0

Watkins (1998). This explains the need of both resource owners and their customers to apply appropriate strategies for hedging price risks.

6.3 Optimal Resource Extraction: Social Welfare View

The Hotelling price trajectory was derived from the profit-maximizing behavior of individual firms. This gives rise to the question of whether the outcome of individual decisions coincides with the interest of society. The model presented below (following Dasgupta and Heal 1979) rests on several assumptions, such as constant population, fixed homogenous reserve stock S, and constant social rate of time preference r (see Sect. 3.4).

Compared to the Hotelling formulation, the model is extended in two ways. First, the objective is not profit but utility derived from consumption of goods and services (the neoclassical definition of welfare). Second, the exhaustible reserve is not consumed directly but constitutes a factor of production that is used in the making of consumption goods. This second assumption is crucial because it opens up the possibility of substitution between exhaustible resources and reproducible capital, which can grow over time. In continuous time with an infinite time horizon, the objective function is the discounted sum of utilities W derived from consumption C_t ($t = 0, 1, 2, \ldots, \infty$),

$$W = \int_0^\infty U(C_t)\, e^{-rt} dt. \tag{6.17}$$

Using r to discount future utilities implies attributing reduced weight to the utility of future generations. This is a debatable assumption. However, without discounting, the infinite integral (6.17) would not converge toward a finite value.

There are two constraints that must be observed. First, the stock of the reserve S is finite, causing the sum of extractions $R_t \geq 0$ to be limited,

$$\int_0^\infty R_t \; dt \leq S. \tag{6.18}$$

Since it would not make economic sense to leave reserves in the ground at the end of an infinitely long planning horizon, this constraint becomes an equality. Moreover, (6.18) masks an equation of motion reflecting the effect of reserve extraction on the value of the remaining stock.[4] This becomes evident when the remaining stock S_t is differentiated with respect to time t,

[4]If the reserve is renewable as e.g. wood, a different formulation is appropriate: The decrease of reserve depends on extractions from and additions to the reserve through regeneration processes, which often are a function of the remaining stock.

$$R_t = -\frac{dS_t}{dt}. \tag{6.19}$$

The second constraint relates consumption C_t to output Q. Output is given by a production function with capital K_t and currently extracted reserves R_t as its arguments. If the production function is independent of time (thus neglecting technological change), consumption is given by

$$C_t = Q(K_t, R_t) - \frac{dK_t}{dt} - \delta K_t - cR_t \tag{6.20}$$

Net investment dK_t/dt, depreciation δK_t, and unit cost of extraction cR_t must be deducted from output $Q(K_t, R_t)$ to obtain the quantity available for consumption. In principle, the production function should also have labor as an argument (see Sect. 5.3), but in view of the assumed constancy of population and technology (implying constant labor productivity), neglecting labor does not modify the core findings while simplifying the analysis.

The objective function (6.17) cannot be analyzed using the methods of static optimization because the constraints, represented by the equations of motion, tie stock and flow variables together. This calls for dynamic optimization methods (see Dasgupta and Heal 1979; Chiang 1992). The point of departure is a generalized Lagrange function,

$$\begin{aligned} L = \int_0^\infty U(C_t)\, e^{-rt} dt + \lambda_t \left(R_t - \frac{dS_t}{dt} \right) + \\ \omega_t \left(Q(K_t, R_t) - C_t - \frac{dK_t}{dt} - \delta K_t - cR_t \right). \end{aligned} \tag{6.21}$$

The first term is the original objective function. The second term takes the constraint (6.19) into account, using the Lagrange multiplier λ_t. This multiplier indicates the extent to which the value of the objective function would diminish if in period t one unit of the reserve were to be extracted in excess of the optimal change in the stock dSt/dt. Therefore, λ_t reflects the opportunity cost of the reserve, which changes over time (in contradistinction to the Hotelling formulation, where it is constant). The third term introduces the constraint on consumption, using the Lagrange multiplier ω_t. This multiplier indicates the extent to which the value of the objective function would diminish if consumption in t were to grow in excess of optimal consumption. Therefore, ω_t denotes the opportunity cost of consumption, which also may vary over time.[5]

In the language of dynamic optimization theory, the Lagrange function has two state variables, K_t and S_t and two control variables, C_t and R_t. As the model has an infinite time horizon, no conditions concerning the final state need to be imposed.

---

[5]The term ‘shadow price’ is sometimes used instead of ‘opportunity cost’.

For the derivation of first-order optimality conditions, the so-called Hamiltonian function is defined. One obtains two equivalent formulations,

$$\begin{aligned} &H = U(C_t)\,e^{-rt} + \lambda_t\left(R_t - \frac{dS_t}{dt}\right) + \omega_t\left(Q(K_t,R_t) - C_t - \frac{dK_t}{dt} - \delta K_t - cR_t\right) \\ &(\text{values discounted to present value}) \\ &H_c = U(C_t) + \lambda_{c,t}\left(R_t - \frac{dS_t}{dt}\right) + \omega_{c,t}\left(Q(K_t,R_t) - C_t - \frac{dK_t}{dt} - \delta K_t - cR_t\right) \\ &(\text{current values}) \\ &\text{with}\,\lambda_t = e^{-rt}\lambda_{c,t}\,\text{and}\;\;\omega_t = e^{-rt}\omega_{c,t}. \end{aligned} \tag{6.22}$$

Whereas the first version refers to discounted utilities and Lagrange multipliers, the second version (to be expounded below) refers to utilities in current values.

For the Hamiltonian function in current values, the first-order optimality conditions read

$$\frac{\partial H_c}{\partial C_t} = \frac{\partial U}{\partial C_t} - \omega_{c,t} = 0,\,\text{implying}\,\omega_{c,t} = \frac{\partial U}{\partial C_t}, \tag{6.23}$$

$$\frac{\partial H_c}{\partial R_t} = -\lambda_{c,t} - \omega_{c,t}\left(\frac{\partial Q_t}{\partial R_t} - c\right) = 0,\;\;\text{implying}\;\;\lambda_{c,t} = \omega_{c,t}\left(\frac{\partial Q}{\partial R_t} - c\right). \tag{6.24}$$

The first condition thus states that the marginal value of consumption must always be equal to its opportunity cost. The second optimality condition shows that additional extraction R_t of the resource has two effects that must balance in the optimum. The first effect is a reduction in the reserve remaining which is valued using the opportunity cost of the reserve $\lambda_{c,t}$. The other effect is the extra production (net of extraction cost), enabling consumption to grow, which is valued using the opportunity cost of consumption $\omega_{c,t}$.

The Hamiltonian function needs to be differentiated also with respect to the two state variables, reserve stock S_t and reproducible capital K_t. They are to be considered as indirect decision variables because they are linked to R_t and C_t through Eqs. (6.20) and (6.21), respectively. Moreover, they affect their pertinent Lagrange multipliers $\lambda_{c,t}$ and $\omega_{c,t}$. The respective optimality conditions read (for the mathematical derivation see Chiang 1992),

$$\frac{\partial H_c}{\partial S_t} = \frac{d\lambda_{c,t}}{dt} - r\,\lambda_{c,t} = 0,\,\text{implying}\;\;\frac{d\lambda_{c,t}}{dt} = r\,\lambda_{c,t}, \tag{6.25}$$

$$\frac{\partial H_c}{\partial K_t} = \frac{d\omega_{c,t}}{dt} - r\,\omega_{c,t} = 0,\;\;\text{implying}\;\;\frac{d\omega_{c,t}}{dt} = r\omega_{c,t}. \tag{6.26}$$

Note that the first condition takes into account that the rate of depletion dS/dt does not depend on the current stock S_t, while the second condition is based on the simplifying assumption that any extra investment and hence acceleration in the buildup of capital, $d/dt(dK/dt) = d^2K/dt^2$, offsets the rate of depreciation δ.

Now Eq. (6.25) implies the so-called Hotelling rule (dropping the t subscript for simplicity),

$$\frac{d\lambda_c}{dt} \cdot \frac{1}{\lambda_{c,t}} = r. \tag{6.27}$$

Therefore, the welfare optimum requires the opportunity cost of the reserve to increase with the rate of social time preference r. If r equals the market interest rate i (which holds if capital markets are perfect), the Hotelling rule coincides with the Hotelling price trajectory. In this case, individual decisions of resource-extracting firms are in accordance with the social welfare optimum.

6.3.1 The Optimal Consumption Path

From Eq. (6.26) in combination with the Hamiltonian function (6.24), the so-called Ramsey rule for the optimal consumption path can be derived. Substitution of (6.24) solved for $d\omega_c/dt$ into (6.26) yields (dropping the t subscript again for simplicity),

$$\frac{d\omega_c}{dt} = -\frac{\partial H_c}{\partial K} + r\omega_c = r\omega_c - \omega_c \frac{\partial Q}{\partial K} = \omega_c (r - Q_K) \text{ with } Q_K := \frac{\partial Q}{\partial K}. \tag{6.28}$$

However, Eq. (6.23) also yields an expression for the opportunity cost of consumption ω_c. Differentiating with respect to time and again simplifying notation gives

$$\frac{d\omega_c}{dt} = \frac{\partial^2 U}{\partial C^2} \cdot \frac{\partial C}{\partial t} \text{ or } \frac{d\omega_c}{dt} = U''(C) \cdot \frac{dC}{dt}, \text{ respectively.} \tag{6.29}$$

Equality of these two equations leads to

$$\omega_c(r - Q_K) = \frac{dC}{dt} U''(C). \tag{6.30}$$

Finally, ω_c can be replaced by the marginal utility of consumption in view of Eq. (6.24)

$$U'(C)(r - Q_K) = \frac{dC}{dt} U''(C). \tag{6.31}$$

Division by C leads to the following expression for the relative change of consumption,

$$\frac{dC}{dt}\frac{1}{C} = \frac{U'(C)(r - Q_K)}{U''(C)C} = \frac{1}{\frac{U''(C)C}{U'(C)}}(r - Q_K). \tag{6.32}$$

The quotient on the right-hand side can be rewritten because the relative change in the marginal utility of consumption $\partial U'(C)/U'(C)$ divided by the relative change of consumption $\partial C/C$ is the same as the elasticity of the marginal utility of consumption with respect to consumption itself (recall that an elasticity relates two relative changes to each other). Under the usual assumption of a decreasing marginal utility of consumption ($U''(C) < 0$), the pertinent elasticity η is defined in a way as to obtain a positive value,

$$\eta := -\frac{\partial[U'(C)]/U'(C)}{\partial C/C} = -\frac{\partial[U'(C)]}{\partial C}\cdot\frac{C}{U'(C)} = -\frac{U''(C)C}{U'(C)} > 0. \tag{6.33}$$

Substitution into Eq. (6.32) leads to the Ramsey rule for the optimal consumption path,

$$\frac{dC}{dt}\frac{1}{C} = \frac{Q_K - r}{\eta}. \tag{6.34}$$

This rule states that consumption must decrease over time unless the marginal productivity of capital Q_K is equal to or exceeds the social rate of time preference r. Given a production function $Q(\cdot)$ without technological change, the marginal product of reproducible capital must decrease when capital stock K grows. This is the law of diminishing marginal returns, which holds when one or more of the other inputs are held constant. In the present case, the other input is the non-renewable resource which is not only held constant but even tends to decline over time. Therefore $Q_K < r$ holds sooner or later, suggesting that consumption C must decrease in the long run. However, a long-run fall in consumption is not only due to the depletion of reserves but also to the discounting of future utilities with $r \geq 0$. In Sect. 6.4.2 below, it is shown that a different formulation of the objective function permits, under certain conditions, a non-declining level of consumption in the long run.

Another element of the Ramsey consumption rule concerns η, the elasticity of the marginal utility of consumption with respect to consumption. This parameter determines the optimal speed of adjustment. Consider a reduction of consumption: Given usual assumptions, this causes the marginal utility of consumption to increase. If the value of η is large, this marginal utility increases fast (the utility function is strongly convex from below), indicating a high loss of utility if consumption is to fall. According to Eq. (6.34), optimal adjustment should be slow in this case. Conversely, if η is small, the utility function is almost linear; therefore, a fall in consumption causes a small loss of utility, indicating that optimal adjustment can be fast.

6.3.2 The Optimal Depletion Path of the Reserve

The model permits to derive another Ramsey rule, this time prescribing the optimal path of reserve depletion. The point of departure is Eq. (6.24), which relates the opportunity cost of the reserve $\lambda_{c,t}$ to the marginal product of the resource Q_R, repeated here for convenience,

$$\frac{\partial H_c}{\partial R_t} = -\lambda_{c,t} - \omega_{c,t}\left(\frac{\partial Q_t}{\partial R_t} - c\right) = 0, \text{ implying } \lambda_{c,t} = \omega_{c,t}\left(\frac{\partial Q}{\partial R_t} - c\right). \quad (6.24)$$

Differentiating with respect to time yields, noting the constancy of unit extraction cost c and dropping the t subscript again,

$$\frac{d\lambda_c}{dt} = \frac{d\omega_c}{dt} Q_R + \omega_c \frac{dQ_R}{dt}. \quad (6.35)$$

Since $d(Q_R - c)/dt = dQ_R/dt$ and in view of Eq. (6.24), the relative change in the opportunity cost of the reserve is given by

$$\frac{d\lambda_c}{dt}\frac{1}{\lambda_c} = \frac{\frac{d\omega_c}{dt} Q_R + \omega_c \frac{dQ_R}{dt}}{\omega_c Q_R} = \frac{d\omega_c}{dt}\frac{1}{\omega_c} + \frac{dQ_R}{dt}\frac{1}{Q_R}. \quad (6.36)$$

In the optimum, the relative change in the opportunity cost of the reserve must therefore be equal to the sum of relative changes in two other parameters,

- The marginal utility of consumption;
- The marginal productivity of the reserve.

According to the Hotelling rule, the rate of change in λ_c needs to be equal to the social rate of time preference r for overall optimality, implying $(d\lambda/dt)/\lambda = r$. Moreover, Eq. (6.28) can be divided by the opportunity cost of consumption ω_c to obtain $d\omega_c/\omega_c = r - Q_k$. Substitution of these expressions into the left-hand and right-hand sides of Eq. (6.36) yields

$$r = r - Q_K + \frac{dQ_R}{dt}\frac{1}{Q_R}. \quad (6.37)$$

From this, the Ramsey rule for the optimal resource depletion path follows immediately,

$$Q_K = \frac{dQ_R/dt}{Q_R}. \quad (6.38)$$

Optimally, the marginal product of the exhaustible resource Q_R must therefore increase at a rate equal to the marginal product of capital Q_K. In the absence of technological change, the marginal product of a production factor can only be increased by using less of it (law of diminishing marginal returns). Thus, the increase of the marginal product of the resource over time implies a diminishing rate of extraction.

6.3.3 Causes and Implications of Market Failure

Given equality of the social rate of time preference and the market rate of interest, society's welfare is maximized if the Hotelling rule is satisfied. Therefore, it would be beneficial for this rule to govern resource markets. As shown in Sect. 6.2, this would also be the optimal solution for reserve-extracting firms, at least under idealistic assumptions such as perfect information and atomistic competition. Yet the reality of resource markets is usually far from satisfying these assumptions. Besides discrepancies between the rate of social time preference and the market rate of interest caused by capital market imperfections, variations in the cost and quality of reserve deposits, and the presence of external costs, there are three problems that are of particular relevance in the context of exhaustible energy sources.

1. Suboptimal extraction capacity: For the Hotelling rule to hold, the rate of extraction R_t needs to be sufficiently flexible. However, geological conditions, limited extraction capacities and bottlenecks in logistics can cause R_t to fall short of the value required by the Hotelling rule. To keep extraction (and with it, production) on the optimal trajectory, additional investment in resource extraction is needed as a rule, which however may create excess capacities given that R_t is to decline in the near future. Anticipating this, companies tend to underinvest in extraction relative to the level that would be necessary to satisfy the Hotelling rule.

   The decision concerning investment in extraction capacity is based on a calculation at the margin. If extraction capacity falls short of the value required by the Hotelling rule, society suffers economic losses. Yet, a capacity expansion according to this rule may also cause losses due to future underuse of this capacity. In the optimum, the two losses must be equal. As long as capacity is below its Hotelling value, extraction occurs at the rate compatible with maximum capacity utilization. This causes the price of the resource to be in excess of the Hotelling price path, with the discrepancy indicating the opportunity cost of the capacity bottleneck. However, this discrepancy decreases over time since depletion of the reserve drives the resource price up and the extraction rate down until the bottleneck no longer exists. The market price then catches up with the Hotelling price path, as shown on the left-hand side of Fig. 6.5 (note that the abscissa is not time but cumulated extraction). Thus, the deviation from the optimal trajectory is transitory rather than permanent in this case.

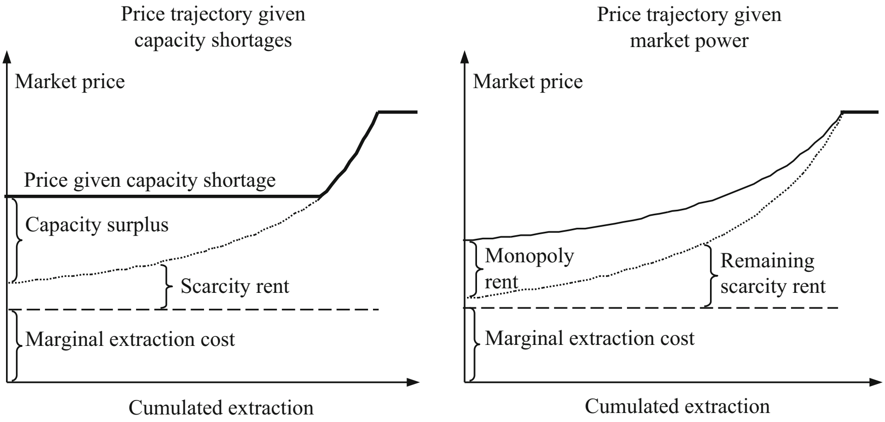

Fig. 6.5 Prices in the presence of capacity shortages and market power

2. Exercise of market power by cartels and monopolies: Market power also causes a deviation from the Hotelling price path. To the extent that owners of extraction rights succeed in imposing a price in excess of the competitive level, demand for the resource falls short of the volume predicated by the Hotelling path. This causes depletion to be slower than under competitive conditions. Therefore, the scarcity rent λ_c is lower than in the competitive case (see the right-hand side of Fig. 6.5), again noting that the abscissa shows cumulated extraction). However, due to the monopoly, the market price of the resource is higher than given competition.

   On the one hand, one may hail the slowing of reserve depletion and the concomitant mitigation of environmental effects ("the monopolist is the environment's best friend"). On the other hand, society's welfare suffers (unambiguously in the absence of external costs) because of the higher resource price and the associated loss of consumer surplus. This time, violation of the Hotelling rule continues as long as prices are affected by monopoly power.
3. Market rate of interest higher than the social rate of time preference ($i > r$): A discrepancy of this type may be due to capital market imperfections. A high interest rate implies a high extraction rate, causing time to depletion to be shortened. The rate of interest may be high because financiers demand a surcharge for the risks associated with exploration, which they tend to overestimate due to information asymmetries. Another reason are poorly defined property rights. This may occur if several firms extract from the same deposit while reserves are geologically mobile, as is the case with conventional crude oil and natural gas fields. This creates an incentive for each company to extract as much of the resource as possible, to the detriment of its competitors (this is known as the common pool problem). An excessive rate of extraction is also to be expected if companies fear expropriation of their rights (through so-called nationalization). Finally, excess extraction can occur in situations where the

holders of the rights (who decide about the extraction rate) are distinct from the owners of the reserve (who claim the scarcity rent). In this case, the scarcity rent fails to provide an economic signal concerning the rate of extraction, which is necessary for the Hotelling price rule to work.

6.4 Sustainability

The sustainability concept originated in forestry, meaning that the use of timber corresponds to the maximum harvest that is compatible with a constant stock of trees. However, contrary to the timber industry, energy sources such as crude oil, natural gas, coal, and uranium are not renewable. The sustainability concept thus cannot be applied without modification to non-renewable energy sources since the stock of the reserves is constant only when extraction is abandoned altogether.

A definition of sustainability that is more suitable to non-renewable energy sources has been proposed by the Brundtland report. According to that report, any development is sustainable "(. . .) that meets the needs of the present without compromising the ability of future generations to meet their own needs" (WCED 1987, p. 43). Accordingly, a decreasing stock of non-renewable energy sources is admissible as long as the needs of future generations can be met with a reduced availability or even lack of non-renewable energy. This definition is called weak sustainability in distinction to strong sustainability, which calls for always keeping a minimum stock of reserves in favor of future generations.

6.4.1 Potential of Renewable Energy Sources

Weak sustainability is only conceivable if there is a sufficient potential of renewable energy sources globally, amounting to multiples of present global use of primary energy. In fact, the potential of renewable energy sources such as hydropower, solar radiation, wind, biomass, ocean energy, and geothermal energy is abundant. The energy of solar radiation hitting the outer atmosphere amounts to 0.14 W/cm^2 (the so-called solar constant). The insolated surface of the Earth is given by

$$6366^2 \pi \left(\mathrm{km}^2\right) = 1\,273 \times 10^{15} \left(\mathrm{cm}^2\right) \tag{6.39}$$

This corresponds to an energy inflow of 178,000 TWa, of which the continents receive 25,000 TWa. Some 6% of total radiation energy hits deserts and wastelands that have no alternate land use. If that solar energy could be transformed into usable energy with an energy efficiency of only 10%, the world would dispose of 37.5 TWa or 28 bn toe, respectively (the technical potential), which is a multiple of today's global energy consumption of about 12.7 bn toe (see Table 2.5).

Table 6.4 Worldwide potential of renewable energy sources

	Theoretical potential (EJ/a)	Technical potential (EJ/a)	Used potential 2013 (EJ)
Biomass (incl. non commercial energy)	2200	160–270	50.0
Hydropower	200	50–60	25.1
Geothermal energy	1500	810–1545	2.3
Wind energy	110,000	1250–2250	4.2
Ocean energy	1,000,000	3240–10,500	–
Solar radiation	3,900,000	62,000–280,000	0.8
Primary energy share			13.5%

EJ= 1 Exajoule = 10^{18} J = 2.39 bn toe
Sources: GEA (2012) and BP (2014)

Therefore, if used to its potential, solar radiation alone would be sufficient to fully eliminate the use of non-renewable energy sources worldwide (see Table 6.4). Additional potential energy sources come from wind, biomass, and possibly nuclear fusion. The technologies required for their use are known in principle. Non-renewable energy sources are thus entirely substitutable as long as there are no other, non-energetic constraints (such as scarcity of precious metals and rare earths as necessary inputs), and as long as harvest rates (see Sect. 2.4) in excess of 1 can be attained.

The limited use of renewable energies is mainly due to their still rather high cost, which in turn is caused by their low energy density (defined as energy flow per m^2 surface or m^3 volume). Therefore capacities for collecting renewable energy require relatively high volumes of material and capital. A second cause, related in particular to solar, wind, and tidal energy, is their discontinuous availability that usually implies low rates of capacity utilization and a backup system in the case of renewable electricity (see Sect. 12.2). Yet, there have been substantial preindustrial uses of renewable energy sources such as biomass and hydropower, benefiting from the fact that nature offers collectors for free in the guise of woody plants and rivers. Accordingly, harvesting these sources of energy is relatively cheap. Indeed, biomass and hydropower continue to constitute the most important renewable sources of the global energy system.

6.4.2 Hartwick Rule for Weak Sustainability

As shown in Sect. 6.3, the Hotelling rule implies a price signal that incentivizes the efficient extraction and use of non-renewable reserves. However, the corresponding Ramsey consumption trajectory [see Eq. (6.34)] does not ensure an increasing or at least non-declining level of consumption over time. Indeed, future generations could be confronted with a drop in their consumption possibilities, violating the criterion of weak sustainability (see Fig. 6.6).

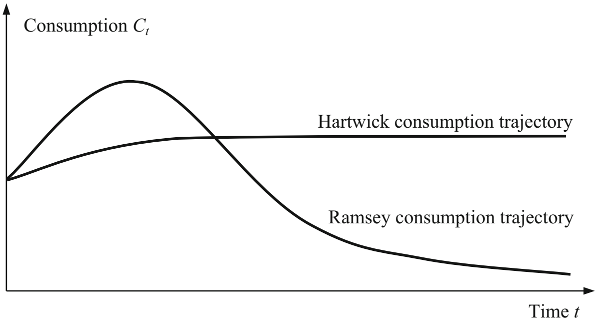

Fig. 6.6 Ramsey and Hartwick consumption trajectories

Going beyond neoclassical welfare optimization, economist Solow (1986) introduced the postulate that per-capita consumption must not decrease, in keeping with weak sustainability. Given a constant population, the Solow postulate reads

$$\frac{dC_t}{dt} \geq 0 \quad \text{for all} \quad t. \tag{6.40}$$

Loosely speaking, this guarantees that the welfare of future generations is at least as high as that of the present generation. Additional consumption by the present generation is admissible only if all future generations can attain at least the same or a higher level of consumption. This defines the so-called Hartwick consumption trajectory (Hartwick 1977, see Fig. 6.6). It is derived as follows.

In order to have a positive rate of consumption in spite of exhaustible energy sources such as crude oil and natural gas, it must be possible to substitute these sources completely with renewables at some future time. For achieving this, resource input R needs to be replaced by reproducible (also called manmade) capital K. An example of complete substitution are plants for hydrogen electrolysis that produce alternative fuels using wind power and photovoltaics. If complete substitution of this type can be attained, a positive amount of consumption should in principle be possible in all future periods.

Whether or not the current level of consumption (per capita) is sustainable in the future crucially depends on the answers to two questions:

– How easily can non-renewable reserves be substituted with reproducible capital?
– Is the current generation willing to finance the necessary growth of reproducible capital by partly renouncing to current consumption?

The first question refers to the elasticity of substitution σ_{RK} between the exhaustible resource R and reproducible capital K. As explained in Sect. 5.3.3, the elasticity of substitution indicates how much the cost-minimizing mix of factor inputs (in the present case R and K) adjusts to a change of relative factor prices (unit price of the resource p_R and cost of capital p_K) given that output is to be kept constant,

$$\sigma_{RK} = -\frac{d(R/K)/(R/K)}{d(p_K/p_R)/(p_K/p_R)} = -\frac{d\ln(R/K)}{d\ln(p_K/p_R)}. \tag{6.41}$$

Here, $\sigma_{RK} \geq 1$ indicates easy substitutability; if $\sigma_{RK} < 1$, the exhaustible resource and capital are not easily substitutable. In that case, the predicted increase of p_R (relative to p_K) requires a disproportionately high increase of capital in order to keep the level of production (and with it, consumption) constant, presumably rendering complete substitution impossible.

Assuming $\sigma_{RK} \geq 1$, thus substitutability between reserves R and capital K, the second question has still to be addressed. Here, the Hartwick rule states that weak sustainability ($dC/dt = 0$) is achievable provided the scarcity rent associated with the resource is entirely invested in reproducible capital. The scarcity rent is given by $(p_R - c)\cdot R$, i.e. the excess of the resource price over the unit extraction cost multiplied by the quantity of the resource used in production. Thus the Hartwick rule can be written

$$\begin{aligned} &\frac{dK}{dt} = (p_R - c)\cdot R \text{ and in its differentiated form,} \\ &\frac{d^2K}{dt^2} = d\left(\frac{(p_R - c)\cdot R}{dt}\right). \end{aligned} \tag{6.42}$$

The proof that this rule ensures weak sustainability given $\sigma_{RK} \geq 1$ proceeds as follows. First, the production function $Q_t = Q(K_t, R_t)$ is differentiated with respect to time,

$$\frac{dQ}{dt} = Q_K\frac{dK}{dt} + Q_R\frac{dR}{dt} \quad \text{where} \quad Q_K := \frac{dQ}{dK}, Q_R := \frac{dQ}{dR}. \tag{6.43}$$

Therefore, Q_K and Q_R denote the marginal productivities of capital and exhaustible resources, as before. An increase in the capital stock contributes to output depending on its marginal productivity Q_K, while an increase in resource input contributes to output depending on its marginal productivity Q_R.

Production can be used for consumption, gross investment (net investment dK/dt plus depreciation $\delta\cdot K$), and for recovery of the cost of extraction $c\cdot R$ [see Eq. (6.20)],

$$Q(K,R) = C + \frac{dK}{dt} + \delta K + cR. \tag{6.44}$$

Differentiation with respect to time leads to

$$\frac{dQ}{dt} = \frac{dC}{dt} + \frac{d\left(\frac{dK}{dt} + \delta K + cR\right)}{dt}. \tag{6.45}$$

At this point, the Hotelling price path is invoked. It states that the scarcity rent $(Q_R - c)$ per unit resource must grow in step with the real rate of interest, which in turn equals the marginal productivity of capital [see Eq. (6.25)]. Therefore, one has

$$Q_K = \frac{\frac{d(Q_R - c)}{dt}}{Q_R - c} = \frac{\frac{dQ_R}{dt}}{Q_R - c}. \tag{6.46}$$

The second equality sign takes into account that $dc/dt = 0$ since unit extraction cost is constant by assumption (for a relaxation of assumptions in several dimensions as well as a critical interpretation of the Hartwick rule, see Mitra et al. 2013). Substitution of Eq. (6.46) into Eq. (6.43) yields

$$\frac{dQ}{dt} = \frac{\frac{dQ_R}{dt}}{Q_R - c}\frac{dK}{dt} + Q_R\frac{dR}{dt}. \tag{6.47}$$

To show sufficiency for achieving a non-declining consumption path $dC/dt \geq 0$, the Hartwick rule is assumed to be satisfied.[6] This means that Eq. (6.42) can be used to replace dK/dt in Eq. (6.47), resulting in

$$\frac{dQ}{dt} = \frac{\frac{dQ_R}{dt}}{Q_R - c}(Q_R - c)R + Q_R\frac{dR}{dt} = R\frac{dQ_R}{dt} + Q_R\frac{dR}{dt}. \tag{6.48}$$

Provided the Hartwick rule holds, the change of aggregate production can thus be reduced to the sum of two terms:

- The change in the marginal productivity of the resource, weighted by the quantity of the resource;
- The change in resource use, weighted by its marginal productivity.

Equation (6.48) is the result of the differentiation of a product. Therefore, one has

$$\frac{dQ}{dt} = \frac{d(Q_R R)}{dt}. \tag{6.49}$$

Finally, solving eq. (6.45) for dC/dt using (6.49) and rearranging terms results in

[6] Withagen and Asheim (1998) have shown that the Hartwick rule is also a necessary condition for weak sustainability.

$$\frac{dC}{dt} = \frac{d(Q_R R)}{dt} - \frac{d\left(\frac{dK}{dt} + \delta K + cR\right)}{dt} = \frac{d(Q_R R - cR)}{dt} - \frac{d^2 K}{dt^2} - \frac{d\delta K}{dt} \\ = \frac{d(p_R - c)R}{dt} - \frac{d(p_R - c)R}{dt} - \frac{d\delta K}{dt} = 0 \tag{6.50}$$

if the amount of capital depreciation δK is a constant. The last equality sign uses the differentiated form of the Hartwick rule (6.42). As a consequence, consumption remains constant over time ($dC/dt = 0$) as long as this rule is satisfied. If the present generation complies with it, future generations will be able to enjoy the same level of consumption as today's population.

However, the question remains whether this steady consumption level is strictly positive ($C > 0$) or not. Solow (1974) provided an answer using the Cobb-Douglas production function

$$Q = \alpha K^{\beta} R^{\gamma} \quad \text{with} \ \ \alpha, \beta, \gamma > 0. \tag{6.51}$$

According to Solow, maximum possible consumption in this case is given by[7]

$$C_{\max} = (1 - c)(S_0(\beta - \gamma))^{\frac{\gamma}{1-\gamma}} K_0^{\frac{\beta-\gamma}{1-\gamma}} \ \text{with initial values} \ K_0 \ \text{and} \ S_0. \tag{6.52}$$

Eq. (6.52) shows that a sustainable positive consumption level $C_{max} > 0$ requires two conditions to be simultaneously satisfied.

- $c < 1$: The marginal cost of extraction must not exceed the marginal productivity of the resource; otherwise, the reserve S_0 would lower output to begin with, causing the buildup of capital to be counter-productive (this is intuitive because the two factors of production are used in combination).
- $\beta > \gamma$: The elasticity of output with respect to the capital input must exceed the elasticity of output with respect to the resource input (which is intuitive, too).

The Cobb-Douglas production function is characterized by a unitary elasticity of substitution between the two inputs ($\sigma_{RK} = 1$). If $\sigma_{RK} > 1$, a higher level of consumption than the one determined by Eq. (6.52) can be sustained, whereas if $\sigma_{RK} < 1$, a positive level of consumption is impossible in the long run because production without any use of the exhaustible resource cannot be attained. Figure 6.7 illustrates the two cases. In the left-hand panel, a production function with an elasticity of substitution $\sigma_{RK} = 0.75$ is shown. No output is possible given $R = 0$ (at point A_0, for example). The right-hand panel shows a production function

[7]Given a constant rate of social time preference r, this (flat) consumption path is not optimal in the sense of the objective function (6.17) of Sect. 6.3. However, it can be shown that if r decreases over time according to $r_t = (1+a \cdot t)^{-b}$ with $a, b > 0$, then the sustainable level of consumption according to Eq. (6.52) is also optimal.

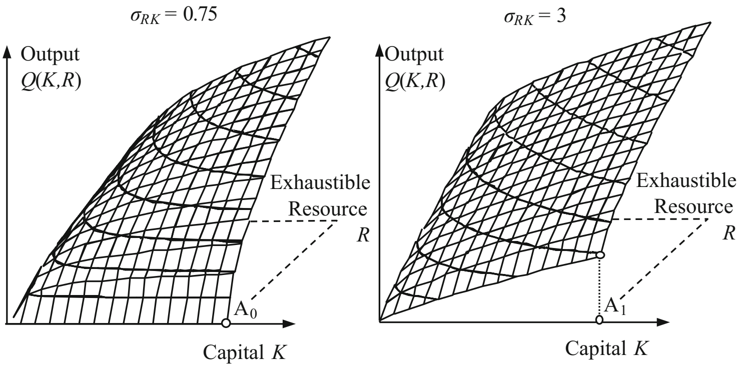

Fig. 6.7 Production function with alternative elasticities of substitution

with $\sigma_{RK} = 3$. Here, positive output ($Q > 0$) is feasible given $R = 0$ (at point A_1, for example).

Evidently, the elasticity of substitution is crucial for weak sustainability. Yet it is a local property of the production function and may change its value when inputs of capital K_t and reserves R_t vary in the course of time t. Indeed, the increasing scarcity of the exhaustible resource causes the two inputs to change over time. Weak sustainability requires that the average value of the substitution elasticity along its trajectory $\{\sigma_{RK,t},\ t{=}t_0,\ t_1,\ldots\}$ is larger or equal to one, permitting the non-renewable resource to be completely substituted by reproducible capital.

The fact that the elasticity of substitution is a local property of the production function has important implications. This can best be explained using the example of wind power, which constitutes an option for replacing fossil fuels. With existing technologies, wind power can already today replace some fossil fuels.[8] However, the best sites for wind generation are the first to be occupied, leaving inferior sites for additional investment designed to substitute fossil fuels. Without technological change, future substitution may therefore become more difficult (σ_{RK} decreases). In this case, the expansion of wind power generation may bind a great deal of capital in the long run, leaving less production output for consumption. This may jeopardize the weak sustainability condition $dC/dt \geq 0$ even though substitutability between fossil fuels and wind power obtains at present.

6.4.3 Population Growth and Technological Change

Section 6.4.2 defines weak sustainability to mean non-decreasing per-capita consumption. Given a production function without technological change and an elasticity of substitution $\sigma_{RK} = 1$, this condition can only be satisfied if population does

[8]Substitution could be based on charging electric vehicles using wind power or on electrolysis which uses wind power to produce hydrogen (known as power-to-gas technology).

not grow. However, even with a growing population weak sustainability can be attained if—in addition to stocking up reproducible capital according to the Hartwick rule—factor productivities increase faster than population due to increased know-how and technological change.

Focusing on the latter, the simplest modeling approach is to view technological changes as exogenous and to incorporate it in the Cobb-Douglas production function (6.51) (see Stiglitz 1974),

$$Q_t = \alpha \cdot K_t^{\beta} \cdot R_t^{\gamma} \cdot e^{f\,t} \text{ with parameters } \alpha, \beta, \gamma, f > 0. \tag{6.53}$$

If the rate of technological change f exceeds the rate of population growth, satisfaction of the Hartwick rule ensures weak sustainability, i.e. non-decreasing per-capita consumption.

The weakness of this approach is that it takes technological change as exogenous. In reality, it is endogenous, driven by (costly) investment in research and development. The many alternatives of modeling this endogeneity cannot be discussed here (see Stoneman 1983 for a survey). Suffice it to remark that 'knowledge' (the stock of know-how and human capital) could be introduced as a factor of production of its own. Its special feature is that it does not decrease but rather increases thanks to learning by doing, contrary to natural resources (through extraction) and reproducible capital (through depreciation). Moreover, its rate of increase depends positively on the amount of know-how already accumulated.

Following up on this idea, total capital stock can be viewed as consisting of reserves of exhaustible resources, reproducible capital, and human capital. In keeping with the Hartwick rule, the scarcity rent derived from resource extraction must entirely be invested in reproducible capital. However, growth in knowledge and know-how in excess of population growth permits to increase the rate of production and consumption per capita. Thereby future generations can attain a higher per-capita consumption level than present generations in spite of an increasing scarcity of exhaustible resources.

6.4.4 Is the Hartwick Rule Satisfied?

The Hartwick rule for weak sustainability demands that the scarcity rent from mining and extracting exhaustible resources be entirely invested as reproducible capital rather than used for consumption purposes. This requirement motivated the countries bordering on the North Sea to abstain from paying the revenues from their oil and gas fields into their social security schemes (i.e. for current consumption) but to rather devote them to investment.

Pearce and Atkinson (1998) checked the extent to which the Hartwick rule may be satisfied by resource-extracting countries. The authors define the total stock of capital as the sum of reserves of natural resources and reproducible capital. For this total stock not to decrease, aggregate savings must exceed net revenue from resource extraction $(pR - c) \cdot R$ plus depreciation of reproducible capital δK,

$$Savings \geq (p_R - c)R + \delta K. \tag{6.54}$$

Countries with a high rate of savings[9] and little reserve extraction—among them Japan and many countries of Western Europe—turn out to be on the path of (weak) sustainability, at least during the observation period. Brazil, Indonesia, the United Kingdom, and the United States are borderline cases because their rate of savings (which is relatively low) combines with a good deal of reserve extraction. However, the African countries sampled fail to satisfy the Hartwick rule. Yet Proops et al. (1999) showed that this assessment changes drastically as soon as international trade in resources is accounted for. In particular, oil-exporting as well as oil-importing countries were found to live off their future generations. By way of contrast, Weitzman (1997) estimated the United States to be in accord with the rule; due to technological change, its future production and consumption possibilities increase much faster than they diminish due to resource extraction.

Political implementation of the Hartwick rule is an issue of its own. Norway is one of the first countries to follow it. Aware of the fact that the country's oil and gas reserves in the North Sea are limited, the Norwegian government began in 1990 to transfer its revenues from oil and gas sales to the Norwegian Government Pension Fund, which is not part of the public budget but is administered by the Norwegian Central Bank. The assets of the Fund, being invested on the international capital market, are exposed to the volatility of stock prices. Their use is decided by the Norwegian parliament, who has credited only the returns (adjusted for inflation) to the public purse until today, leaving the principal intact. With roughly 130 bn EUR at the end of 2004 and 600 bn EUR by the end of 2014, the Fund is one of the largest sovereign wealth funds worldwide. While its later use is not decided yet, the Hartwick rule suggests long-term investments in the development of alternative energy sources, infrastructure, education, and research. However, in view of fast growth of its assets, there is considerable political pressure to use it for consumption purposes as well.

The Norwegian Fund is designed not only to implement the Hartwick rule but also to insulate the public budget from the volatilities of oil and gas prices, and to protect the economy from the so-called Dutch disease. The Dutch disease is a scenario which can occur in small countries with an important resource extraction sector. The large-scale expansion of this sector generates important export revenues which usually are exchanged in domestic currency. This demand drives up the domestic currency, causing domestic goods to become expensive compared to foreign goods. As a consequence, the country's international competitiveness suffers, hampering its exports of other goods and services (e.g. by fisheries in the case of Norway).[10]

[9]Macroeconomic savings divided by the Gross Domestic Product.

[10]This phenomenon was first observed in The Netherlands at the beginning of an export boom at the beginning of the 1970s, hence its name 'Dutch disease'.

Indeed, large oil and gas deposits may turn out to be a curse rather than a blessing for many economies. Norway is one of the few energy-exporting countries to have clearly benefited up to now, motivating several oil-exporting countries to copy Norway by creating similar sovereign wealth funds.

Yet the accumulation of the scarcity rents derived from the extraction of oil and gas can pose another problem if it results in high amounts relative to global capital markets. According to the Hotelling price trajectory, scarcity rents grow over time, while according to the Hartwick rule, they need to be invested rather than consumed. However, do global capital markets offer sufficient investment opportunities? What happens to the (real) interest rate when the supply of capital continues to increase? What if the funds are invested in financial instruments only, in response to a lack of productive investment opportunities? Is the global financial system stable at all? Indeed, international capital markets may not be capable of accommodating the global inflow of scarcity rents which has surpassed 1000 bn USD annually, equivalent to 2% of the world's Gross Domestic Product. In addition, concentration of these funds in the hands of a few oil-exporting countries poses a particular risk to the countries hosting this foreign investment.

References

Adelman, M. A. (1990). Mineral depletion, with special reference to petroleum. *The Review of Economics and Statistics, 72*(1), 1–10.

BGR. (2014). *Reserven, Ressourcen und Verfügbarkeit von Energierohstoffen (Reserves, resources, and availability of energy Resources)*. Hannover: Bundesanstalt für Geowissenschaften und Rohstoffe. Retrieved from www.bgr.bund.de/DE/Themen/Energie/Downloads/Energiestudie_2014.pdf

BP. (2014). *BP statistical review of world energy*. Retrieved from www.bp.com/statisticalreview/

Campbell, C. J., & Laherrere, J. H. (1996). *The world's oil supply 1930–2050*. Geneva: Petroconsultants.

Chiang, A. C. (1992). *Elements of dynamic optimization* (pp. 210–211). New York: McGraw-Hill.

Dasgupta, P. S., & Heal, G. M. (1979). *Economic theory and exhaustible resources*. Cambridge: Cambridge University Press.

Erdmann, G., & Zweifel, P. (2008). *Energieökonomik - Theorie und Anwendungen (Energy economics – theory and applications)* (2nd ed.). Berlin: Springer.

GEA. (2012). *Global energy assessment – toward a sustainable future*. Cambridge University Press. Retrieved from www.iiasa.ac.at/web/home/research/Flagship-Projects/Global-Energy-Assessment/Home-GEA.en.html

Hartwick, J. M. (1977). International equity and the investing of rents from exhaustible resources. *American Economic Review, 67*(5), 972–974.

Hotelling, H. (1931). The economics of exhaustible resources. *Journal of Political Economy, 39* (2), 137.

Hubbert, M. K. (1956). Nuclear energy and fossil fuels. In American Petroleum Institute (Ed.), *Drilling and production practice*. New York: American Petroleum Institute.

Hubbert, M. K. (1962). *Energy resources. A report to the Committee on Natural Resources*. Washington, DC: Government Printing Office. Publication No. 1000-D.

IEA. (2000). *Experience curves for energy technology policy*. Paris: International Energy Agency.

Kaufmann, R. K., & Cleveland, C. J. (2001). Oil production in the lower 48 states: Economic, geological, and institutional determinants. *The Energy Journal, 33*(1), 27–49.

Mitra, T., et al. (2013). Characterizing the sustainability problem in an exhaustible resource model. *Journal of Economic Theory, 148*, 2164–2182.

Nakicenovic, N. (2002). Technological change and diffusion as a learning process. In A. Grubler, M. Nakicenovic, & W. Nordhaus (Eds.), *Technological change and the environment* (pp. 160–181). Washington, DC: Resources for the Future Press.

Pearce, D., & Atkinson, G. (1998). The concept of sustainable development: An evaluation ten years after Brundtland. *Swiss Journal of Economics and Statistics, 134*(3), 251–269.

Plourde, A., & Watkins, G. C. (1998). Crude oil prices between 1985 and 1994: How volatile in relation to other commodities? *Resource and Energy Economics, 20*, 245–262.

Proops, J., Atkinson, G., Scholtheim, B., & Simon, S. (1999). International trade and the sustainability footprint: A practical criterion for its assessment. *Ecological Economics, 28* (1), 75–98.

Reynolds, D. B. (2002). Using non-time-series to determine supply elasticity: How far do prices change the Hubbert curve? *OPEC Review, June*, 149–167.

Solow, R. M. (1974). Intergenerational equity and exhaustible resources. *Review of Economic Studies, 41* (*Symposium on the Economics of Exhaustible Resources*), 29–45.

Solow, R. M. (1986). On the intergenerational allocation of natural resources. *Scandinavian Journal of Economics, 88*, 141–149.

Stiglitz, J. E. (1974). Incentives and risk sharing in sharecropping. *Review of Economic Studies, 41* (*Symposium on the Economics of Exhaustible Resources*), 219–256.

Stoneman, P. (1983). *The economic analysis of technological change*. Oxford: Oxford University Press.

WCED. (1987). *From one earth to one world: An overview*. Oxford: World Commission on Environment and Development.

Weitzman, M. L. (1997). Sustainability and technical progress. *Scandinavian Journal of Economics, 99*(1), 1–13.

Withagen, C., & Asheim, G. (1998). Characterizing SUSTAINABILITY: The Converse Hartwick's rule. *Journal of Economic Dynamics and Control, 23*, 159–165.