doi: 10.1111/j.1467-6419.2008.00549.x

SHOULD WE WORRY ABOUT THE FAILURE OF THE HOTELLING RULE?

Tobias Kronenberg

Forschungszentrum Jülich, Institute of Energy Research – Systems Analysis and Technology Evaluation (IEF-STE)

Abstract. The continuing dependence of the global economy on fossil fuels is worrying because it imposes limits on growth due to the non-renewable nature of these resources and also contributes to global climate change. Resource optimists believe that this is no reason to worry, because the economy will always find a way to overcome these constraints. Their arguments, however, require that resource prices reflect the scarcity of non-renewable resources, which implies that they must obey the *Hotelling rule*. Empirical analyses, however, show that the Hotelling rule does not hold in reality, which raises the question: does the failure of the Hotelling rule imply that social optimality is not achieved? This paper argues that the answer depends on the reason for the failure. If extraction and exploration costs, or technological progress in these activities, are the reasons for the failure, a market failure is not implied, and optimality may still be achieved. But if the Hotelling rule fails due to uncertain property rights or strategic interaction, the market will surely fail to provide an optimal solution. A market failure is likely to speed up resource consumption compared to the social optimum.

Keywords. Climate change; Hotelling rule; Natural resources; Non-renewable resources

1. Introduction

Throughout human history, economic growth has been accompanied by increasing energy use (Rifkin, 2002). In recent decades, the link between energy consumption and economic activity appears to have been weakened, as GDP has grown faster than energy consumption, which some economists interpret as the result of a ‘dematerialization’ of GDP. If this trend of increasing dematerialization were to continue, it could result in a ‘de-linking’ of GDP from energy consumption, which would enable the former to grow whereas the latter would fall (Smulders, 1995). However, there are limits to dematerialization. Van Zon and Yetkiner (2003) point out that there are certain minimum input requirements because ‘people cannot live in virtual houses and live on virtual food’. Stern and Cleveland (2004) argue that the increase in energy efficiency was mainly due to a shift from poor- to high-quality fuels. Because this shift is already largely complete, they conclude that ‘prospects for further large reductions in energy intensity seem limited’. Thus, it seems safe to

Journal of Economic Surveys (2008) Vol. 22, No. 4, pp. 774–793

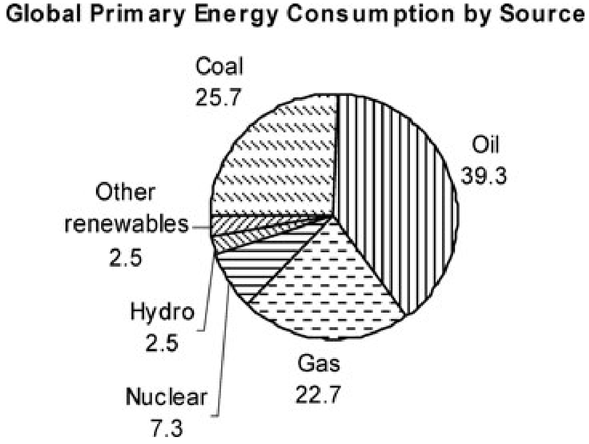

Figure 1. Breakdown of Energy Consumption in 2000.
Source: *World Energy Outlook* 2002 and author's calculations.

conclude that the global economy will continue to be dependent on a large amount of energy input for the foreseeable future.

Figure 1 shows that in the year 2000 oil accounted for the lion's share of almost 40% of global primary energy consumption. The two next most important energy sources were coal (25.7%) and natural gas (22.7%). Hence, fossil fuels in total accounted for 87.7% of total primary energy. The remaining energy came from nuclear sources (7.3%), hydropower (2.5%) and other renewables (also 2.5%).

The dominant position of fossil fuels in the global energy mix is problematic because the burning of these fuels causes a negative externality in the form of carbon dioxide emissions, which contribute to global warming. This negative externality is not taken into account by individual polluters, and therefore the total level of emissions is too high. Furthermore, as Sinclair (1994) shows, in a dynamic context a further complication arises. If society discounts the future, it is optimal to postpone the negative effects of emissions by shifting fossil fuel consumption into the future, but because of the externality problem individual agents fail to do this, which is clearly inefficient. Thus, the negative stock externality of carbon dioxide emissions implies that social welfare can be improved by delaying the consumption of fossil fuels.

Another cause for concern is the non-renewable nature of fossil fuel resources. The dependence on non-renewable resources is often regarded as a huge problem, because when these resources are used up energy consumption patterns will have to be adjusted at a great cost. On the other hand, the finiteness of fossil fuels can be regarded as a blessing, for as fossil fuel consumption must fall, carbon dioxide emissions must also eventually fall. This does not mean that mankind must literally 'run out of oil' some day. Rather, the literature on resource economics suggests that the increasing scarcity of fossil fuels will be reflected by rising prices according to the *Hotelling rule* (Hotelling, 1931). In its most simple form, which is based on very restrictive assumptions such as perfect information and costless extraction, it states that the price of a non-renewable resource should grow at the real rate of interest (Lassere, 1991). This simple form, however, is clearly refuted by the empirical evidence. Resource economists such as Pindyck (1978), Farzin (1984) and André and Smulders (2004) have developed more complex versions of the Hotelling rule, attempting to improve its empirical validity by adding more realistic assumptions. Others have come to firmly reject the Hotelling approach. Banks (2004), for instance, insists that the *Hubbert curve* approach, which is named after the geographer M. King Hubbert, who used it to predict with some precision the peak in US oil production, is more appropriate.

The validity of the Hotelling rule in either form is of central importance to the climate change issue. Resource optimists, who believe that the economy will always overcome the constraints imposed by the finite ecosystem of the Earth, support their view with different arguments including demand shifts, increasing exploration and recycling, replacement of natural resources by man-made capital and technological progress (Neumayer, 2000). In this optimistic view, the link between economic growth and increasing carbon dioxide emissions will diminish and, in the long run, be reversed, so that further economic growth will be accompanied by decreasing pollution. This relationship, the so-called 'environmental Kuznets curve', has indeed been observed for some specific pollutants (Grossman and Krueger, 1993, 1995). Whether an environmental Kuznets curve exists for carbon dioxide emissions is an open question (Dijkgraaf and Vollebergh, 2005). If it exists, it may eventually provide a natural solution to the problem of global warming.

The arguments of the resource optimists, however, rely on the assumption that market prices correctly reflect the scarcity of non-renewable resources, which is only the case if they obey the Hotelling rule. In a perfectly competitive market, resource prices do obey the Hotelling rule, thus producing the socially optimal outcome (Weinstein and Zeckhauser, 1975). Consequently, if the Hotelling rule does not hold there must be some market failure, and hence the time path of resource consumption is not optimal. On the other hand, we know that the stock externality of carbon dioxide emissions already distorts the time path of resource consumption. Thus, there are at least two market failures affecting the consumption of fossil fuel resources. We know that one of them, the stock externality, increases the speed of resource consumption. In principle, it might be possible for the failure of the Hotelling rule to cause another market failure which slows down resource consumption, and, in the end, the time path of resource consumption may be quite close to the optimum. Therefore, we need to answer two questions: does the failure of the Hotelling rule imply an (additional) market failure, and if yes, does it increase or decrease the speed of resource consumption?

This paper contributes to the literature by evaluating the competing explanations for the 'Hotelling failure' and examining their impact on the market's ability to allocate the consumption of fossil fuels over time in an efficient manner. To this end, the following section describes the basic model of resource depletion based on the work by Hotelling (Hotelling, 1931). Section 3 confronts this model with the empirical evidence, showing that the model's predictions appear to fail miserably. Section 4 provides traditional explanations for the failure, which are mainly based on technological and geological factors. Section 5 presents two new explanations based on institutional factors. Section 6 analyses the effects of these failures on social welfare and proposes some policy-relevant implications, and Section 7 sets forth a conclusion.

2. Non-Renewable Resources in Theory

In its most simple version, the Hotelling rule states that the price of a non-renewable resource should rise at the rate of interest. Although resource economists have known for a long time that this simple version of the Hotelling rule is clearly refuted by empirical evidence, it remains in frequent use because of its simplicity and intuitive appeal. Some researchers, for example Antony (2007) and Conrad (2004), use it as a shortcut to pin down the development of resource prices in models which focus on other aspects of the economy. These models offer valuable insights into the issues on which they focus, but the use of the simple Hotelling rule is not as innocuous as it may seem. Perman *et al.* (2003) argue that although the empirical failure of the Hotelling rule does not make it false as a theory, it does mean that the rule 'only applies to the idealized world for which it was constructed'. Consequently, the same applies to all models which use the Hotelling rule as a shortcut. Such models are still valuable, but it must be made perfectly clear that because they do not 'fit the facts' of the real world, they cannot be used to explain or predict real world phenomena. They can only be used to explain phenomena that occur in the 'idealized world' of the Hotelling rule.

It is therefore worthwhile to take a closer look at this 'idealized world'. In this section, a simple version of the Hotelling model of resource depletion is presented to illustrate the intuition that underlies the Hotelling rule. It is shown that this simple model is based on strong assumptions such as perfect information and extraction at no cost, which are not very realistic. Under these circumstances, it should come as no surprise that the simple Hotelling rule does not fit the facts which are observed in reality.

In the 'idealized world', a certain stock R of a non-renewable natural resource is owned by a profit-maximizing firm, which extracts resources and sells them on a perfectly competitive market. Because the firm is a price taker, it cannot influence the price of the resource, so its only choice variable is extraction in each period t, x_t. Furthermore, extraction costs are assumed to equal zero, so the price p_t of the resource is equal to the marginal profit of extracting a unit of R. The total value of the firm is equal to the sum of its discounted profit flows:

$$V = \sum_{t=0}^{T} p_t x_t (1+r)^{-t} \tag{1}$$

T, which may approach infinity, is the firm's planning horizon. The firm, acting rationally, maximizes its value. It has to respect the constraint that the sum of resource extraction in all periods cannot exceed the total resource stock:

$$\sum_{t=0}^{T} x_t \leq R \tag{2}$$

The firm solves the problem of maximizing its objective function (1) with respect to the constraint (2). The solution to this maximization problem must be found from the following Kuhn–Tucker conditions:

$$\frac{\partial L}{\partial x_t} \leq 0 \quad x_t \geq 0 \quad \text{and} \quad x_t \frac{\partial L}{\partial x_t} = 0$$

Resource extraction x_t cannot be negative, so the second constraint is always fulfilled. The third constraint can only be fulfilled if either x_t is zero (no extraction takes place in period t) or $\partial L/\partial x_t = 0$. Thus, extraction will only take place in period t if $\partial L/\partial x_t = 0$ is fulfilled. Let us assume now that extraction occurs in two consecutive periods t and $t + 1$. This requires

$$(1 + r)^{-t} p_t + \lambda = 0 \tag{3a}$$

and

$$(1 + r)^{-(t+1)} p_{t+1} + \lambda = 0 \tag{3b}$$

Combining these two equations and rearranging terms yields

$$(1 + r) p_t = p_{t+1} \tag{4}$$

Thus, extraction will only occur in two consecutive periods if equation (4) holds. This equation is known as the *Hotelling rule*. It states that the price of the finite resource must grow exponentially at a rate which is equal to the rate of interest to ensure consecutive resource extraction. Hence, the resource price p_t can be written as a function of the initial price p_0 and the interest rate r:

$$p_t = p_0 \mathrm{e}^{rt} \tag{5}$$

The economic interpretation of the Hotelling rule is quite intuitive, because the stock of the natural resource can be regarded just like any other stock of capital. The holders of natural resource capital earn interest in the form of increasing resource prices. Whereas stock market capital earns an interest rate of r, natural capital yields an interest rate of $p_{t+1}/p_t - 1$. By extracting resources in the current period and investing the revenues on the stock market, the firm can freely convert its natural resource assets into other capital assets. If stock market capital were to offer a higher interest than natural capital ($r > p_{t+1}/p_t - 1$), the firm would prefer to convert all its natural resources into stock market capital. If all natural resource holders do the same, this amounts to an increase of supply. Assuming a normal downward-sloping demand curve, the current price p_t must fall. On the other hand, if natural resource capital offered a higher interest than that on the stock market, capital owners would be willing to sell their stock market assets to acquire natural resource assets. Effectively, the demand for natural resources would rise, and this would drive up the price of the resource.

Thus, in a competitive market, any deviations from the Hotelling rule would offer arbitrage opportunities. By taking advantage of these opportunities, investors would affect the price of resources in such a way that it moved closer to the level required by the Hotelling rule, which represents the only possible equilibrium. Hence, the Hotelling rule is not only the solution to the maximization problem of a resource-owning firm but a fundamental equilibrium condition, stating that the price development of finite resources depends crucially on the characteristics of the stock market, i.e. the prevailing (long-term) interest rate.

The Hotelling rule does not only make predictions about the price development of a finite resource; it can also be used to say something about the extraction rate of the resource and the effect that resource scarcity has on growth. Let us consider an economy that produces its output according to the following production function:

$$Y_t = A_t K_t^{\alpha} L_t^{\beta} R_t^{1-\alpha-\beta} \tag{6}$$

R_t denotes the flow of resources in period t, K_t is a stock of accumulated production factors, L_t is the labour force, A_t is a technology shift parameter, and Y_t is output. Under Cobb–Douglas production with constant returns to scale, the expenditure share of each input is constant:

$$\frac{p_t R_t}{Y_t} = 1 - \alpha - \beta = \text{const} \tag{7}$$

If the Hotelling rule holds, p_t grows at a constant rate. This implies that R_t/Y_t, the resource intensity of the economy, falls over time.

In per capita terms (assuming full employment), the production function (6) can be written as

$$y_t = A_t K_t^{\alpha} L_t^{\beta-1} R_t^{1-\alpha-\beta} \tag{8}$$

where y_t denotes output per capita (Y/L). Let us assume that a social planner intends to provide a constant population with a constant level of GDP per capita. What is necessary to fulfil this objective? To answer this question, we first rewrite equation (6) in terms of growth rates (which we denote with a circumflex):

$$\hat{y}_t = \hat{A}_t + \alpha\hat{K}_t + (1-\beta)\hat{L}_t + (1-\alpha-\beta)\hat{R}_t \tag{9}$$

The condition for constant GDP per capita ($\hat{y} = 0$) under zero population growth ($\hat{L} = 0$) is

$$\hat{A}_t + \alpha\hat{K}_t = -(1-\alpha-\beta)\hat{R}_t \tag{10}$$

The right-hand side of equation (10) constitutes the 'resource drag' that is caused by the finiteness of natural resources. From equation (7) we know that under constant output the flow of resource inputs will fall at the same rate as their prices rise, and the Hotelling rule tells us that this rate is equal to the interest rate. Thus, the right-hand side is the product of the expenditure share of resources in output and the interest rate. For output to remain constant, the left-hand side must be equal to the resource drag. Economically, this means that technical progress (growth in A) and factor accumulation (growth in K) can in principle offset the resource drag. Despite the finiteness of resources, GDP per capita can remain at a constant positive value and may even increase if the left-hand side of equation (10) exceeds the right-hand side.

Using rules of thumb, we can get an idea of the relative magnitude of the resource drag. Because $\hat{R}$ is equal to the (negative of the) real interest rate, it seems reasonable to assume a value between 4%, as suggested by Stiglitz (1997, p. 213), and 6%, as suggested by Jones (2002, p. 179), so we settle for a compromise of 5%. The capital share α is probably close to 0.3 if only physical capital is considered and close to 0.6 if human capital is also included (Mankiw *et al.*, 1992). The labour share is then either close to 0.3 or 0.6, depending on the definition of capital. In either case, the resource share $(1 - \alpha - \beta)$ is close to 0.1, so the right-hand side would be 0.1 times 0.05, which is equal to 0.005. With respect to the left-hand side, the growth accounting literature suggests that in the long run total factor productivity (TFP) growth ($\hat{A}$ in this model) is between 1% and 2% (Jones, 2002). Thus, even without any capital accumulation ($\hat{K} = 0$), the left-hand side would be between 0.01 and 0.02, which is between two and four times larger than the resource drag. These rule-of-thumb calculations suggest that although the finiteness of resources imposes a drag on growth, it is possible to provide a constant population with a constant level of GDP. In fact, it is possible to achieve a growing level of GDP per capita because the left-hand side of equation (10) is larger than the right-hand side, and even if there were no technological progress, sustained growth would still be feasible, because a sufficiently rapid accumulation of man-made capital could overcome the growth drag caused by the finiteness of non-renewable resources.

Another important property of the Hotelling rule is that it constitutes not only the optimal solution for the resource owner's profit maximization problem; it also provides the optimal solution for society as a whole. This should not be surprising, because the rule was derived under optimistic assumptions such as perfect information, perfect competition and absence of externalities. Using optimal control methods, a social planner would arrive at exactly the same solution as the private resource owner (Kuuluvainen and Tahvonen, 1995). These considerations might lead to the impression that although the finiteness of natural resources imposes a drag on growth, the market mechanism will ensure that these resources are allocated optimally in time, so that growth as a whole is optimal. Technical progress and capital accumulation then stand a good chance of overcoming the resource drag and generating long-term economic growth despite the finiteness of natural resources, as argued above. Solow (1974) shows that growth may be sustainable even without technological change as long as the income share of man-made capital is larger than that of natural capital, and Hartwick (1977) derives an optimal savings rule which states that an economy should invest the entire resource rent in new capital. The Hotelling rule together with the *Hartwick rule* has since become the cornerstone of resource economics.

The above line of reasoning, however, is based on a number of restrictive assumptions, in particular the assumption that man-made capital can be substituted for natural resource inputs. This assumption is a major point of disagreement between neoclassical economists, who argue that substitution between the two factors is relatively easy, and ecological economists, who argue that this substitution is difficult if not impossible because non-reproducible natural resources are embodied in 'man-made' capital. Due to this difference, ecological economists tend to be much more sceptical about the feasibility of sustainable economic growth (Cleveland and Ruth, 1997).

Keeping these important caveats in mind, we will now take a look at the data on natural resource prices to see whether the Hotelling rule provides a good approximation to reality.

3. Non-Renewable Resources in Reality

The model presented in the previous section is a very simplistic one, but well into the 1970s it was considered state of the art. This section shows the failure of the simple Hotelling rule when confronted with empirical data. Section 4 will then present a number of subsequent extensions to the basic Hotelling model, including extraction costs, exploration, technological change and imperfect competition.

Figure 2 shows how prices of four major industrial resources evolved during the 20th century. All prices are adjusted for inflation and indexed with 1949 as the base year to allow for graphical comparability. For the most important resource, crude oil, we only have information from 1949 on. For the three other resources, namely copper, zinc and iron ore, we have information from 1900 on.

Crude oil prices show a slight downward trend from 1949 to the early 1970s. The two oil crises of the 1970s are strikingly visible. From 1973 to 1974, the real price of crude oil rose by more than 60%, reached a plateau, and then from 1978 to 1981 it almost tripled. It then declined during the 1980s and 1990s to reach a

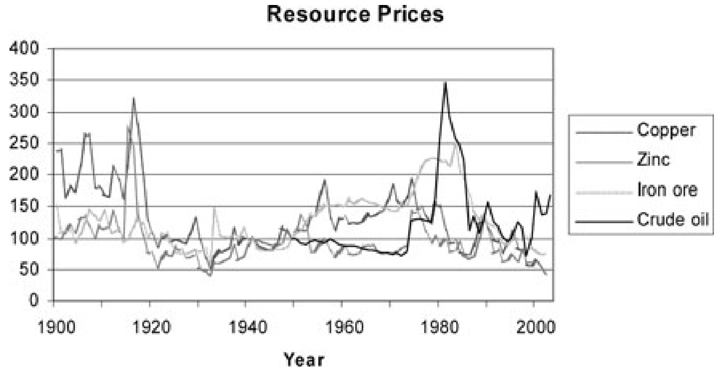

Figure 2. Real Price Development of Four Major Natural Resources.

Source: websites of the Energy Information Administration and the US Geological Survey.

low in 1998, when crude oil was cheaper (in real terms) than before the first oil shock. From 1998 to 2003 its real price increased again.

Although it would be outside the scope of this paper to provide a detailed analysis of the development of oil prices during the past few decades, we can nevertheless summarize the main observations. First, oil prices have become much more volatile after 1974. They followed a smooth and stable decline from 1949 to 1973, but after that they have been greatly influenced by all sorts of economic and political crises. The OPEC oil embargo against a number of Western countries, the Iranian revolution, and the Gulf War of 1991 are examples of crises which led to soaring oil prices. Other crises, such as the global recession of the early 1980s and the Asian crisis of 1997, caused a pronounced fall in oil prices. Second, oil prices have generally not increased in real terms since 1949. During most of the time period a trend of falling oil prices is obvious, and despite the two oil shocks, crude oil was actually cheaper in 1998 than in 1949 (in real terms). From 1998 to 2003 there was a marked increase in oil prices, which was partly due to the fast growth of China's huge economy and the associated increase in oil demand. In general, however, oil prices after 1973 have been volatile but not increasing.

The prices of the other resources exhibit similar features. Over the 20th century as a whole, there was a downward trend in resource prices. From 1900 to 2000, the prices fell by almost 40% in the case of zinc and copper, and by 20% in the case of iron ore. There were times of increasing prices, too. The real price of iron ore, for instance, doubled between 1945 and 1964, and zinc and copper became much more expensive from 1973 to 1974, thus echoing the oil shock. In general, however, the prices of these vital natural resources fell more often than they rose. An econometric analysis by Ocampo and Parra (2003) concludes that 'relative raw materials prices deteriorated markedly in the course of the twentieth century'.

4. Reconciling the Theory with the Evidence

The previous section showed that the simple Hotelling rule is at odds with empirical observations. The simple rule, however, was derived under very restrictive assumptions. In this section, we relax a few of these assumptions and examine whether the resulting theoretical implications confirm more closely to reality.

4.1 *Costly Extraction*

The simple Hotelling rule was derived under the assumption that resources can be extracted at zero cost. This may be an acceptable approximation in the case of oil-abundant countries like Saudi Arabia, but not in the North Sea or Alaska, where the costs of oil drilling are clearly not zero. Therefore, it may be more reasonable to allow for positive marginal extraction costs.

This extension requires only a slight reformulation of the basic Hotelling model. The Hotelling rule, as expressed in equation (4), states that the marginal profit of extracting in period t should be equal to the discounted marginal profit of extracting in period $t + 1$. Marginal profit is equal to marginal revenue minus marginal cost, so equation (4) can be seen as a special case in which marginal cost equals zero. Thus, we can rewrite it as

$$(1 + r)(p_t - mc) = p_{t+1} - mc \tag{11}$$

This equation may look puzzling at first, because we assume perfect competition, and hence there should be no profits. The difference between resource price and marginal cost, in this case, is not a profit in the economic sense. It is a 'royalty', or the *in situ* value of the resource. The latter term indicates the value of leaving the resource in place (*in situ*) instead of removing it. Expressed in other words, it is the opportunity cost of extracting the resource, because extracting at the present time means that less extraction is possible in the future. Let us simplify the notation by defining $q_t = p_t - mc$ as the *in situ* value of the resource in period t. This allows us to write

$$q_t = q_0 \mathrm{e}^{rt} \tag{12}$$

We can thus see that actually it is not the resource price which grows at the rate of interest, but the *in situ* value of the resource. Under zero marginal extraction cost the two are the same, but with positive marginal extraction cost they are quite different. The above equation allows us to write prices as a function of q_0 and r:

$$p_t = q_0 \mathrm{e}^{rt} + mc \tag{13}$$

Because the resource price is the sum of the royalty, which grows at rate r, and mc, which is constant, it is easy to see that the resource price may grow at a faster or slower rate than the royalty, depending on the development of the marginal extraction cost. For example, one may wish to assume that there is some positive marginal extraction cost. Certainly, such an assumption would be more realistic than the assumption of extraction at no cost, but it still could not explain the falling resource prices that we observe in reality. According to the modified Hotelling rule (equation (11)), resource prices should still grow, albeit at a rate which is lower than the interest rate.

The assumption of constant marginal extraction cost is still unrealistic. Hotelling himself argued in his seminal paper that there may be *stock effects* in the sense that the marginal extraction cost depends negatively on the remaining stock. This may be the case if easily accessible resource deposits are extracted first, and over time more inaccessible deposits have to be tapped. The effect is that the marginal extraction cost rises over time. Because the resource price is the sum of *in situ* value and marginal extraction cost, and the latter does not grow over time, prices must grow faster than under the simple Hotelling rule. Thus, stock effects in the form of rising marginal extraction cost contradict the empirical evidence of constant or falling resource prices to an even greater extent.

It is also possible, however, that these costs may decrease over time due to technological progress. This assumption does not negate the basic insight of the Hotelling model, namely that the *in situ* value of the resource must grow at the rate of interest. The model is still similar to that described above. The resource price is a combination of *in situ* value, which grows over time, and marginal extraction cost, which falls due to technical progress. Thus, there are two opposing effects. The increasing *in situ* value tends to raise the resource price, whereas the falling extraction costs tend to reduce the resource price. If initial extraction costs are high and technological progress is fast, the latter effect may dominate the former, which would lead to a period of falling resource prices. In the long run, however, the *in situ* effect will dominate, and prices will rise again. Technical progress can thus explain a U-shaped price development, where prices at first fall and then rise over time (Krautkraemer, 1998). Slade (1982) finds empirical evidence for such a U-shaped price path for several mineral resources. André and Smulders (2004) present a model with endogenous technological progress in extraction, which also produces a U-shaped price path for non-renewable resources.

If technological change is indeed the reason for the failure of the Hotelling rule, there are two important implications. First, the trend of falling resource prices is a temporary one, because in the long run the exhaustion effect will overcome the cost reduction effect, and resource prices must increase. Second, if we allow for technological change, there may be repercussions in other sectors of the economy. In the model by André and Smulders, for example, the economy increases its efforts to bring down extraction costs at the expense of other R&D efforts. TFP growth falls, and GDP growth as well. The R&D market, however, is characterized by severe market failures, which raises doubts about its ability to provide an optimal mix between improvements in TFP and reductions in extraction cost.

Until further research is done on this subject, only speculation is possible. One might suspect that because the benefits of extraction cost reductions accrue to a small group of agents, namely resource owners, whereas the benefits of TFP growth are spread over all agents in an economy, the former type of technological change may be overprovided by the market. This would constitute another market failure that needs to be addressed. However, in order to draw more reliable conclusions, additional research is required.

4.2 *Exploration*

In the basic Hotelling model, it is assumed that the total stock of the resource is finite and known with certainty. In reality, however, new resource deposits are constantly being discovered by chance or through intentional exploration efforts. Each new discovery increases the known resource stock.

If there are stock effects, exploration can lower the marginal extraction cost by increasing the stock of remaining resources. Assuming that the initial known stock is very small, and exploration increases the known stock by a large amount, the fall in the marginal extraction cost could be quite substantial. In such a case, the resource price may fall in periods which are characterized by successful exploration. In the long run, however, exploration opportunities are limited, and exploration will run into diminishing returns. As new discoveries become less frequent, the basic Hotelling intuition holds again, and resource prices rise again. Thus, allowing for exploration also generates the possibility of a U-shaped price development (Krautkraemer, 1998).

However, the existence of exploration opportunities will not go unnoticed, and will affect agents' expectations. Such expectations will be formed about the frequency and size of new discoveries, and also about the total amount of resources that will be discovered. Exploration then simply becomes a costly activity which can be added to marginal extraction costs, and agents will no longer base their decisions on the known remaining reserves but on the *expected* remaining reserves. Expectations will be revised whenever new information is revealed, which will generate some volatility and deviations from the Hotelling rule. Nevertheless, the basic Hotelling intuition still applies, and the resource price must increase unless expectations are systematically incorrect.

4.3 *Imperfect Competition*

We have seen above that the Hotelling rule in equation (4) states that the present value of marginal profit in any period should be equal to that of any other period, and that equation (4) is a special case in the sense that marginal extraction costs are zero. It is also a special case in another respect, namely that marginal revenue equals price. This is only the case under perfect competition, however. For a monopolistic supplier, marginal revenue is

$$mr_t = p_t + p'_t x_t \tag{14}$$

The second term on the right-hand side of this equation reflects the fact that an increase in supply x_t lowers the price p_t, and this reduces total revenue. Thus, for a monopolistic supplier, marginal revenue is lower than the price. For our analysis, this implies that a monopolistic resource supplier operates under a modified Hotelling rule. Substituting equation (14) for p_t in equation (4) yields

$$p_{t+1} + p'_{t+1} x_{t+1} = (1 + r)(p_t + p'_t x_t) \tag{15}$$

To gain a more intuitive understanding, let us rewrite this equation in terms of elasticities. Let us define the price elasticity of demand as

$$\varepsilon_t = \frac{\partial x(p_t)}{\partial p_t} \frac{p_t}{x(p_t)} \tag{16}$$

This allows us to write

$$p_{t+1} = (1 + r) p_t \frac{1 + 1/\varepsilon_t}{1 + 1/\varepsilon_{t+1}} \tag{17}$$

Thus, the time path of the resource price now depends on the development of the demand elasticity over time. If this elasticity does not change, the equation reduces to the simple Hotelling rule, which again shows that equation (4) is a special case of equation (17). Under perfect competition the price elasticity is infinity, and this does not change over time. Another possibility is an isoelastic demand curve, which is obtained, for example, if the resource is an input into a Cobb–Douglas production function. Along an isoelastic demand curve, the elasticity of demand is constant, and equation (17) again becomes the simple Hotelling rule.

This finding led Stiglitz (1976) to believe that in the market for a non-renewable resource a monopoly behaves exactly like a perfectly competitive firm, thus causing no welfare loss.

However, this finding does not hold if ε changes over time. Depending on the structure of the demand side, the elasticity may either increase or decrease over time. The growth rate of the resource price is now

$$\frac{p_{t+1}}{p_t} = (1 + r)\frac{1 + 1/\varepsilon_t}{1 + 1/\varepsilon_{t+1}} \tag{18}$$

Depending on the development of ε over time, the resource price can increase more slowly or quickly than under perfect competition. The monopoly will time its supply of resources so as to take advantage of the changing elasticity. When the elasticity is high, supply will be high, taking advantage of high quantity and still reasonable prices. When the elasticity is low, supply will be low, pushing prices to extremely high levels. The elasticity of demand may change even without any changes in technology. The growing resource price implies that over time we move along the demand curve, and if the demand curve is not isoelastic its elasticity differs at each point along the curve. In this case the behaviour of a monopoly differs from that of a competitive firm (Gaudet and Lassere, 1988).

In addition to the simple monopoly model, the literature on resource economics also offers the so-called cartel-versus-fringe model. This model can explain a more complicated price development, with resource prices falling for some amount of time and then rising again. In the long run, however, the resource price must still follow an upward trend. The empirical validity of the cartel-versus-fringe model is very difficult to test because of a lack of data on *in situ* values (Withagen, 1998).

5. Institutional Explanations for the Failure

Traditionally most of the criticism of the Hotelling approach came from natural scientists, mostly geologists. It is therefore not surprising that the first explanations for the Hotelling failure relied on ‘hard-science’ factors, for example the technological difficulties of extracting a resource, the search for new resource deposits, and the progress made in these processes. Institutional factors have so far not featured prominently in the literature on resource economics, except for the contributions dealing with market power, but even these rely ultimately on geological factors, namely the spatial distribution of resource deposits leading to clustering and increasing returns to scale. In the following sections, we propose two new institutional explanations which may contribute to our understanding of the failure of the Hotelling rule.

5.1 *Uncertain Property Rights*

The simple Hotelling rule, as most of its extensions, assumes that property rights over resources are well defined and protected. In reality, however, countries that are richly endowed with natural resources are often plagued by massive corruption and weak institutions. In fact, the corruption associated with natural resources such as oil or minerals is one of the main causes of the ‘curse of natural resources’ (Kronenberg, 2004). The concession rights to exploit these resources are often granted by a government which is (1) corrupt and (2) not democratically elected. The holder of these concession rights must always fear the loss of these rights if either he loses the favour of the government or the government itself is toppled by a revolution or a coup. Assuming perfectly secure property rights may therefore be too optimistic, especially in the long run.

The resource owner will form some expectations about the probability of him losing the concession rights. Let us assume that the probability of retaining control of the resources in the next period is η. Then, the expected present value of the resource revenue becomes

$$V = \sum_{t=0}^{T} p_t x_t \, [(1+r)\eta]^{-t} \tag{19}$$

Because η is smaller than one, the value to the current resource owner is smaller than the social value of the resource. The same optimization procedure as above now yields a modified Hotelling rule:

$$(1+r)\eta p_t = p_{t+1} \tag{20}$$

This equation shows that with uncertain property rights the required growth rate of the resource price is lower. In fact, if η is small enough to fulfil the condition $(1+r)\eta < 1$ the resource price will actually fall over time. As long as η is smaller than one, so that there is some uncertainty about property rights, the time profile of the resource price will be flatter (i.e. rising more slowly and possibly falling) than under certain property rights. This flatter time profile is only possible if the time path of resource extraction is steeper (i.e. falling more quickly). Thus, uncertain property rights can make the Hotelling approach consistent with the observed fall in real resource prices (Figure 2).

Equation (20) shows that under uncertain property rights the resource owner will have an incentive to extract the resource more quickly than the social optimum would require. This result is important, because it implies that uncertain property rights constitute a market failure raising the speed of resource extraction above the social optimum, which is detrimental to welfare. Policy measures that reduce the speed of extraction may then be welfare enhancing.

It would be interesting to know whether uncertain property rights play a significant role in reality. Casual evidence suggests that property rights can be extremely uncertain, and wars have been waged to change the ownership of resources. The historical hostility between France and Germany, for example, was to a large extent fuelled by the rich coal deposits in the much-contested province of Alsace-Lorraine, and the Iraqi invasion of Kuwait in 1990 was also connected with disputes over access to rich oil deposits. It is sometimes believed that the increasing scarcity of fossil fuels may lead to more ‘resource wars’ in the future. On the other hand, the property rights to many other resource deposits have been remarkably stable over centuries. A thorough empirical investigation may shed more light on this issue, but is outside the scope of this study.

5.2 *Strategic Interaction*

Another possible deviation from the Hotelling rule may stem from strategic interaction between the suppliers and consumers of a non-renewable resource. Assume that there are two players in a game, a resource owner O and a resource consumer C. The total stock of resources is known to O, but not to C. O may attempt to communicate this information to C, but C need not believe O, because O may have an incentive to lie.

Without strategic interaction, intertemporal optimization would require extraction to take place at a constant rate which depends on the interest rate, as in equation (4). We now assume that C has the option of developing a backstop technology, which would provide a perfect substitute for the natural resource. This, however, requires costly R&D efforts and takes time. A social planner could maximize social welfare by choosing the optimal timing of resource extraction and R&D.

However, the invention of the backstop technology renders the remaining resource stock worthless. Therefore, O will want to delay the arrival of the backstop technology. But it is C who makes the decision when to develop the technology. O can only delay the arrival of the backstop technology by influencing C's decision. And he can achieve this by lying about the stock of remaining resources, on which C has no information.

Assume that the true stock of resources available at time zero is R_0. The optimal time path of extraction depends positively on R_0. Therefore, if C observes a certain amount of extraction $x(t)$ taking place in period t, he will form expectations about R_0. Naturally, the higher $x(t)$, the higher his expectations about R_0. Thus, by extracting 'too much' in the early stages of the game, O can influence C's expectations of R_0. C will overestimate R_0, and will consequently delay the development of the backstop technology.

On the other hand, C can also attempt to influence O's decisions. If C announces that he will develop a backstop technology very soon, and O happens to believe the announcement, O will make sure he gets rid of the resources before they become worthless, and will increase the extraction rate. This allows C to buy the resources at a lower price.

Thus, if there are information asymmetries between the owners and consumers of a resource, strategic interaction takes place, and credible announcements play a critical role. Specifically, resource owners will have an incentive to overestimate the resource stock, so as to delay the development of substitutes for the resource. To make this announcement credible, they have to follow an extraction path consistent with the overestimated resource stock, so extraction will be faster than socially optimal. Resource consumers will have an incentive to announce the development of a backstop technology, and resource owners will react to this threat by raising the extraction rate and lowering the resource price (if the demand curve is downward sloping). In both cases, resource extraction occurs faster than socially optimal.

Evidence for cases of strategic interaction is difficult to find due to the nature of the problem. Nevertheless, some authors have argued that resource owners, especially in the oil industry, do indeed overestimate their resource reserves on purpose. They certainly have an incentive to do so: for oil companies, their stock price depends on the value of assets, so if they provide larger reserve estimates their stock price will rise. OPEC countries also have an incentive to boost their reserve estimates, because their export quotas depend on the amount of reserves they have. Campbell and Laherrère (1998) suspect that in the 1980s ‘six of the 11 OPEC nations increased their reserve figures by colossal amounts [. . .] only to boost their export quotas’. Such episodic evidence suggests that the figures of global oil (and other resource) reserves are indeed systematically overestimated. This overestimation would lead to a serious market failure, increasing the rate of resource consumption above the optimum and delaying the development of alternative energy technologies. Again, there may be scope for welfare enhancement if governments enact policies that reduce the rate of resource consumption and/or speed up the development of alternative energy technologies.

6. Market Failures and Policy Intervention

The discussion in the previous two sections has shown that the failure of the Hotelling rule does not necessarily imply a failure of the market to allocate resources efficiently over time. If extraction costs and technical progress are the reasons for the failure of the Hotelling rule, there is no problem. The market solution can still be the optimal one, but it may differ from the simple Hotelling rule due to these forces. However, technological change is also associated with much uncertainty, and the market for R&D is subject to serious distortions. It may be necessary to correct these market failures in order to achieve an optimal provision of technological change.

If new discoveries are the reason for the Hotelling failure, there may be a problem if expectations are not fully rational. When new discoveries occur, people may revise their expectations about the total available stock of the resource. This is not a problem if expectations are rational. The market will then price the resource according to the expected total stock, and will adjust to any positive or negative realizations. But if expectations are not fully rational, there is a problem. Under optimistic expectations the rate of resource consumption would be too high, and under pessimistic expectations it would be too low. In theory, therefore, the bias can go towards either direction. Recent empirical evidence suggests that market participants ‘generally believe in a mean-reverting process for the price of oil in the process of bidding and accepting offers for reserves, which means that they expect the price of oil to increase when it is below historical average and to decrease when it is above average’ (Schiozer *et al.*, 2006). As argued above, however, the price of oil must increase in the long run, which implies that expectations about the future price and the availability of oil are overly optimistic, indicating that current oil consumption is too high.

If uncertain property rights or strategic interaction are the reason, there is definitely a problem. A non-renewable resource is valued properly only if the property rights are certain. If there is a small chance of losing control over the resource, possibly due to war or revolution, the value to the owner is ‘smaller than it should be’. A dictator, for example, has control over a country’s natural resources, but there is a certain chance of him being overthrown by a coup or a revolution. This means that he has an incentive to exploit the resource as fast as possible, sell it on the world market, and store the revenue in some safe place (for example a foreign bank account). Resource extraction would then occur too quickly.

Similar problems arise if strategic interaction between resource suppliers and consumers enters the picture. A sequential game between these two parties will not generally lead to the socially optimal outcome. Resource owners have an incentive to overstate their true resource endowment in order to encourage the consumers’ dependence on the specific resource. Resource extraction is then likely to occur more quickly than is optimal.

These findings must be evaluated in combination with the negative externality imposed by carbon dioxide emissions. Because the consumption of fossil fuel resources causes a negative stock externality, such resources will generally be consumed too quickly from a social point of view because of the discounting of the future. Under a positive discount rate, it makes sense to feel pleasure early and postpone pain. Therefore, an omniscient social planner would choose to postpone the pain that results from burning oil (climate change). In reality, however, no omniscient social planner exists. Decisions are made by self-interested producers or profit-maximizing firms who do not take externalities into account. These observations led Sinclair (1994) to the conclusion that the consumption of fossil fuels occurs too quickly. Grimaud and Rouge (2005) come to the same conclusion although they treat pollution as a flow, whereas Sinclair treats it as a stock, and assume that pollution affects utility, whereas Sinclair assumes that it affects productivity. Because the same finding emerges from different sets of assumptions, it seems to be quite robust.

7. Conclusion

To sum up, the simple Hotelling rule does not hold in reality because a number of its assumptions are violated. If the deviations from the simple Hotelling rule are caused by marginal extraction costs, which could rise due to stock effects or fall due to technical progress, this does not pose a serious problem. Under these conditions it is socially optimal to follow a modified Hotelling rule, and that is what the market will do, so the market solution is efficient. If, however, uncertain property rights or strategic interaction are the causes of the failure of the Hotelling rule, this is quite unfortunate because it implies that the market will not lead to an efficient solution.

The question is whether the shortcomings of the market can be overcome by means of policy. In the case of uncertain property rights, the solution is easier said than done: establish well-defined property rights. However, the very lack of property rights suggests that the government itself is either unwilling or unable to solve the problem, because otherwise they would have already established these property rights. And with strategic interaction between consumers and suppliers, a solution is also difficult to find. Because most non-renewable resources like oil and coal are traded on global markets, consumers and suppliers are often located in different countries, so that a national government cannot impose an efficient solution. Information asymmetries and lack of trust between different national governments have been a persistent problem throughout history, and a solution to these problems is not readily at hand. Nevertheless, it is important to acknowledge the presence of these market failures. They suggest that non-renewable resource consumption does indeed occur too quickly, and other policy measures which reduce the resource extraction rate may therefore be welfare enhancing.

In addition to the observations reported by Sinclair (1994) and Grimaud and Rouge (2005), who argue that fossil fuels are consumed too quickly from a social point of view due to the negative externality of global warming, our findings suggest a further inefficiency in the intertemporal allocation of fossil fuel consumption. Rather than offsetting each other, these inefficiencies tend to reinforce each other, suggesting that fossil fuels are indeed consumed too quickly. Correcting these inefficiencies requires a policy intervention at the global level to bring about a lower resource consumption rate.

With respect to the relationship between pollution and growth, the existence of non-renewable resources does not weaken the link. If the Hotelling rule were to hold, resource consumption would fall over time, and if technical progress were indeed fast enough to overcome the ‘growth drag’, the link between growth and increasing pollution would be broken. In reality, however, this has not been the case. The reason for this is the failure of the Hotelling rule. Resource prices have not increased significantly, and consequently resource consumption has not fallen. As long as the Hotelling rule does not hold, one cannot expect a ‘natural’ solution to the growth–environment trade-off.

Acknowledgements

This paper is based on research which was undertaken while the author was employed at the University of Maastricht. The author would like to thank Thomas Ziesemer, Kristin Vetter and Sabine Fuss for helpful comments and advice.

References

André, F.J. and Smulders, S. (2004) Energy use, endogenous technical change and economic growth. In *EEARE 13th Annual Conference*. Budapest.

Antony, J. (2007) Depletion of non-renewable resources and endogenous technical change. Universität Augsburg, Institut für Volkswirtschaftslehre, Volkswirtschaftliche Diskussionsreihe Beitrag 291.

Banks, F.E. (2004) Beautiful and not so beautiful minds: an introductory essay on economic theory and the supply of oil. *OPEC Review* 28: 27–62.

Campbell, C.J. and Laherrére, J. (1998) The end of cheap oil. *Scientific American* 78–83.

Cleveland, C.J. and Ruth, M. (1997) When, where, and by how much do biophysical limits constrain the economic process? A survey of Nicholas Georgescu-Roegen's contribution to ecological economics. *Ecological Economics* 22: 203–223.

Conrad, K. (2004) Network effects, compatibility and the environment: the case of hydrogen powered cars. University of Mannheim, Department of Economics Discussion Papers No. 613-04.

Dijkgraaf, E. and Vollebergh, H. (2005) A test for parameter homogeneity in CO_2 panel EKC estimations. *Environmental and Resource Economics* 32: 229–239.

Farzin, Y.H. (1984) The effect of the discount rate on depletion of exhaustible resources. *Journal of Political Economy* 92: 841–851.

Gaudet, G. and Lassere, P. (1988) On comparing monopoly and competition in exhaustible resource exploitation. *Journal of Environmental Economics and Management* 15: 412–418.

Grimaud, A. and Rouge, L. (2005) Polluting non-renewable resources, innovation and growth: welfare and economic policy. *Resource and Energy Economics* 27: 109–129.

Grossman, G.M. and Krueger, A.B. (1993) Environmental impacts of a North American free trade agreement. In P.M. Garber (ed.), *The Mexico–U.S. Free Trade Agreement*. Cambridge, MA, and London: MIT Press.

Grossman, G.M. and Krueger, A.B. (1995) Economic growth and the environment. *Quarterly Journal of Economics* 110: 353–377.

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

Hotelling, H. (1931) The economics of exhaustible resources. *Journal of Political Economy* 39: 137–175.

Jones, C.I. (2002) *Introduction to Economic Growth*. New York: W. W. Norton.

Krautkraemer, J.A. (1998) Nonrenewable resource scarcity. *Journal of Economic Literature* 36: 2065–2107.

Kronenberg, T.H. (2004) The curse of natural resources in the transition economies. *Economics of Transition* 12: 399–426.

Kuuluvainen, J. and Tahvonen, O. (1995) The economics of natural resource utilization. In H. Folmer, H.L. Gabel and H. Opschoor (eds), *Principles of Environmental and Resource Economics: A Guide for Students and Decision-Makers*. Aldershot: Elgar.

Lassere, P. (1991) *Long-Term Control of Exhaustible Resources*. Chur, Switzerland: Harwood.

Mankiw, N.G., Romer, D. and Weil, D. (1992) A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107: 407–438.

Neumayer, E. (2000) Scarce or abundant? The economics of natural resource availability. *Journal of Economic Surveys* 14: 307–335.

Ocampo, J.A. and Parra, M.A. (2003) The terms of trade for commodities in the twentieth century. *CEPAL Review* 79: 7–35.

Perman, R., Ma, Y., McGilvray, J. and Common, M. (2003) *Natural Resource and Environmental Economics*. Harlow: Pearson.

Pindyck, R.S. (1978) The optimal exploration and production of nonrenewable resources. *Journal of Political Economy* 86: 841–861.

Rifkin, J. (2002) *The Hydrogen Economy*. Cambridge: Polity.

Schiozer, R.F., Costa Lima, G.A. and Suslick, S.B. (2006) The expectations of oil companies on future oil prices: an empirical analysis of reserves trading. Available at SSRN: http://ssrn.com/abstract=942649 (accessed 17 June 2008).

Sinclair, P.J.N. (1994) On the optimum trend of fossil fuel taxation. *Oxford Economic Papers* 46: 869–877.

Slade, M.E. (1982) Trends in natural-resource commodity prices: an analysis of the time domain. *Journal of Environmental Economics and Management* 9: 122–137.

Smulders, S. (1995) Environmental policy and sustainable economic growth: an endogenous growth perspective. *De Economist* 143: 163–195.

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

Stern, D.I. and Cleveland, C.J. (2004) Energy and economic growth. Rensselaer Working Papers in Economics 0410.

Stiglitz, J.E. (1976) Monopoly and the rate of extraction of exhaustible resources. *American Economic Review* 66: 655–661.

Stiglitz, J.E. (1997) *Economics*. New York: W. W. Norton.

Van Zon, A. and Yetkiner, H.I. (2003) An endogenous growth model with embodied energy-saving technical change. *Resource and Energy Economics* 25: 81–103.

Weinstein, M.C. and Zeckhauser, R.J. (1975) The optimal consumption of depletable natural resources. *Quarterly Journal of Economics* 89: 371–392.

Withagen, C. (1998) Untested hypotheses in non-renewable resource economics. *Environmental and Resource Economics* 11: 623–634.