Should we worry about the failure of the Hotelling rule?

Tobias Kronenberg

Despite the recent progress in alternative energy technologies, the global economy remains dependent on fossil energy resources, which is one of the reasons for the climate change problem. The theoretical literature predicts that the prices of these resources should be rising over time by the Hotelling rule, and argues that this behaviour is compatible with social optimality. Rising prices should reflect the increasing scarcity of fossil fuel resources, and the market should react to this by developing alternative energy technologies, thus finding a “natural solution” to the climate change problem.

Empirical analyses show that the Hotelling rule does not hold in reality. The prices of most nonrenewable resource have actually fallen over time. Does the failure of the Hotelling rule also 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, this does not imply a market failure, and optimality may still be achieved. But if the Hotelling rule fails due to uncertain property rights or strategic interaction, the market surely fails, an optimal solution will not be achieved. A market failure is likely to speed up resource consumption compared to the social optimum.

The paper concludes that if the original market failures cannot be mended, a second-best-policy is called for. This policy would attempt to reduce the consumption of fossil fuels and speed up the development of alternative energy technologies.

JEL Classification: Q3, Q5.

Keywords: Hotelling rule, non-renewable resources, climate change

---

1 UNU-MERIT, Maastricht University, Postbus 616, 6200 MD Maastricht, The Netherlands
E-mail: tobias.kronenberg@merit.unimaas.nl
This paper is intended for presentation at the young economists sessions.
1 Introduction

Throughout human history, economic growth has been accompanied by increasing energy use. In recent decades, the link between energy consumption and economic activity appears to have weakened, as GDP has grown faster than energy consumption, which some economists interpret as 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 while the latter falls (Smulders, 1995). However, this trend of dematerialization cannot continue indefinitely. Van Zon and Yetkiner (2003) point out that there are certain minimum energy requirements since “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 quality fuels to high quality fuels. Since this shift is already largely complete, they conclude that “prospects for further large reductions in energy intensity seem limited”.

Although the debate on the de-linking of GDP and energy consumption is still going on, it seems safe to conclude that the global economy will continue to be dependent on a large amount of energy input for the foreseeable future. This energy input comes from different primary sources, as can be seen in Figure 1.

![Global Primary Energy Consumption by Source](image)

Figure 1. Percentage shares of primary energy sources in global energy consumption in 2000. Source: WEO 2002 and author’s calculations
Should we worry about the failure of the Hotelling rule?

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

This dominant position of fossil fuels in the global energy mix is worrying for several reasons. Firstly, the burning of these fuels causes greenhouse gas emissions and thus contributes to global warming. Second, many important deposits of fossil fuel resources are located in politically unstable regions, which may pose a threat to global peace and security. The fossil energy system thus imposes a double negative externality, harming the environment and endangering security. Another cause of concern is the nonrenewable nature of these resources. There is a certain amount of these resources in the ground, and if they are depleted they cannot be renewed. The finiteness of fossil fuels is often regarded as a huge problem, because when we run out of these resources we will have to adjust our energy consumption patterns. On the other hand, the finiteness of fossil fuels can be regarded as a blessing, for as fossil fuel consumption must fall, carbon emissions must also fall. If growth continues despite the lack of fossil fuels, the link between economic growth and increasing pollution would be broken. This would provide a ‘natural solution’ to the global warming problem. However, even if the link is broken through natural forces, it is not clear when this will happen, and whether the timing of resource consumption is optimal.

There is a respectable amount of literature on the optimal consumption of non-renewable resources, based on the work by Harold Hotelling (1931). This literature argues that in the social optimum the price of a non-renewable resource should be rising at the rate of interest, a requirement known as the Hotelling rule (Lasserre, 1991). Furthermore, resource consumption should decline over time, asymptotically falling to zero (Dasgupta and Heal, 1974). Weinstein and Zeckhauser (1975) argue that a competitive market will follow these rules under certain conditions, thus producing the socially optimal outcome. In reality, however, neither the Hotelling rule nor the declining consumption requirement appear to be fulfilled. Some economists have responded to this by firmly rejecting the Hotelling approach in favour of a more realistic one (e.g. Banks, 2004), whereas others – including Pindyck (1978), Farzin (1984), and more recently André and Smulders (2004) – modify it by adding some more realistic features.
The empirical failure of the Hotelling rule is no longer puzzling, and theoretical explanations abound. A more important question, however, is whether the failure of the Hotelling rule implies that the market is failing to produce a socially optimal resource consumption path. This paper contributes to answering this question by evaluating the competing explanations for the “Hotelling failure” and examining their impact on the market’s ability to achieve the optimal outcome. To this end, the following section describes the basic model of resource depletion based on the work by Harold 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 concludes.

2 Nonrenewable Resources in Theory

2.1 Resource Prices

The “Hotelling rule”, based on the work of Harold Hotelling (1931), has become one of the most widely used concepts in environmental economics. It is therefore worthwhile to repeat the derivation and the implications of the Hotelling rule. In this section, a simplified version of the Hotelling model of resource depletion is presented to illustrate the intuition that underlies the Hotelling rule.

Let us assume that there is a certain stock of a natural resource $R$. The resource is nonrenewable and does not depreciate. Therefore, the timing of extraction has no consequences on the total amount that can be extracted. Assuming that the resource is completely extracted over a time horizon of $T$ (which could be a positive finite number or go towards infinity), the sum of extraction in all periods must equal the total resource endowment.

Furthermore we assume that the resource stock $R$ is owned by a profit-maximizing firm, which sells the extracted resources on a perfectly competitive market. Thus, the firm is a price taker. Since it cannot influence the price of the resource, its only choice variable is extraction in period $t$, $x_t$. Let us also assume that extraction costs are 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:
Should we worry about the failure of the Hotelling rule?

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

The firm, acting rationally, maximizes this value. It has to take into account the constraint that the sum of resource extraction in all periods cannot exceed the total resource stock:

(2) \[ \sum_{t=0}^{T} x_t \leq R \]

Combining the objective function and the constraint we can set up the Lagrangian function which must be maximized:

(3) \[ L = \sum_{t=0}^{T} p_t x_t (1 + r)^{-t} + \lambda \left( R - \sum_{t=0}^{T} x_t \right) \]

The Kuhn-Tucker conditions for the solution to this maximization problem are:

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

Resource extraction \( x_t \) obviously 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 \( \frac{\partial L}{\partial x_t} = 0 \). Thus, extraction will only take place in period \( t \) if \( \frac{\partial L}{\partial x_t} = 0 \) is fulfilled. Let us assume now that extraction occurs in two consecutive periods. This requires:

\[ (1 + r)^{-t} p_t + \lambda = 0 \]

and

\[ (1 + r)^{-(t+1)} p_{t+1} + \lambda = 0 \]

We can rewrite these two equations as one:
(1 + r)^{-t} p_t + \lambda = (1 + r)^{-(t+1)} p_{t+1} + \lambda

Rearranging terms yields:

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

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, we can write the resource price $p_t$ as a function of the initial price and the interest rate:

(5) \quad p_t = p_0 e^{rt}

The economic interpretation of this rule is quite intuitive. 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 the stock market capital earns an interest rate $r$, natural capital yields an interest rate of $p_{t+1} / p_t$. 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. Due to arbitrage opportunities, all capital stocks must yield the same interest, so $r$ must be equal to $p_{t+1} / p_t$, which is exactly the relationship described by the Hotelling rule.

### 2.2 Resource Consumption

Let us assume that there is a certain stock of a nonrenewable resource which can be used to produce energy. In period $t$ an amount $E(t)$ of the resource is extracted and transformed into energy. The remaining deposit of the resource is $R(t)$. Mathematically, this can be expressed as:
Should we worry about the failure of the Hotelling rule?

(6) \[ \dot{R}_t = -E_t \]

Assuming a demand function with a price elasticity of one, such as \( E_t^0 = 1/ p_t \), it can be shown that the economy consumes a constant fraction of the remaining resource in each period. Let us call this fraction \( c_R \):

(7) \[ c_R = \frac{E_t}{R_t} \]

Dividing (6) by \( R \), substituting (7) and multiplying by –1 yields:

(8) \[ \frac{\dot{R}}{R} = -c_R \]

Equation (8) tells us that the remaining resource deposit will be falling at a constant rate, and this rate happens to be the consumption rate \( c_R \). Thus, we can write \( R(t) \) as an exponential function of time and the initial resource endowment:

(9) \[ R(t) = R_0 e^{-c_R t} \]

And, since we know from (7) that \( E = c_R R \), we can write \( E(t) \) as:

(10) \[ E(t) = c_R R_0 e^{-c_R t} \]

Thus, since resource extraction is proportional to the remaining deposit, it also falls at a constant rate. This is shown in Figure 2, where both resource extraction \( E \) and remaining deposits \( R \) are falling exponentially over time.
To understand the economic relevance of the Hotelling rule, let us consider what happens on a competitive market if the rule is not currently fulfilled. If stock market capital would offer a higher interest than natural capital \( (r > p_{rt} / p_t) \), 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 would offer 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 rises, and this drives 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 moves closer to the level required by the Hotelling rule. The Hotelling rule thus is the only possible equilibrium. Hence, the Hotelling rule is not only the solution to the maximization problem of a resource-owning firm. It is a fundamental
Should we worry about the failure of the Hotelling rule?

equilibrium condition, stating that the price development of finite resources depends crucially on the characteristics of the stock market, i.e. the prevailing (long-run) 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 of the effect that resource scarcity has on growth. Let us consider an economy that produced its output according to the following production function:

\[ Y_t = A_t K_t^a R_t^{1-a} \]  

This is a simple Cobb-Douglas production function. \( R_t \) is the flow of resources in period \( t \), \( K_t \) is a stock of accumulated production factors (which could be interpreted as physical capital, human capital, or labor 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:

\[ \frac{p_t R_t}{Y_t} = 1 - \alpha = \text{const} \]

If the Hotelling rule holds, \( p_t \) is growing at a constant rate. This implies that the ratio \( R_t / Y_t \) must be falling at the same rate. In words, this means that the resource intensity of the economy is falling over time.

Let us assume that a social planner intends to keep output constant. What is necessary to fulfill this objective? To answer this question, we first rewrite (11) in terms of growth rates (which we denote with a hat):

\[ \dot{Y}_t = \dot{A}_t + a \dot{K}_t + (1 - a) \dot{R}_t \]

This shows us that the growth rate of \( Y \) depends on the growth rates of \( A \), \( K \), and \( R \). The condition for a stable output level is:

\[ \dot{A}_t + a \dot{K}_t = -(1 - a) \dot{R}_t \]
Should we worry about the failure of the Hotelling rule?

The right-hand side of equation (14) constitutes the ‘resource drag’ that is caused by the finiteness of natural resources. From equation (12) we know that under constant output the flow of resource inputs will fall at the same rate as their prices rise, and the Hotelling rules 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 (increase in $A$) and factor accumulation (increase in $K$) can in principle offset the resource drag. Despite the finiteness of resources, output can remain at a constant positive value and may even increase if the left-hand side of (14) exceeds the right-hand side.

Using rules of thumb, we can get an idea about the relative magnitude of the resource drag. Since $\hat{R}$ is equal to the (negative of the) real interest rate, it should be somewhere around five percent, and surely not higher than ten percent. Considering that $K$ is interpreted here as the aggregate of all other production factors, thus including all sorts of reproducible capital, $\alpha$ is probably close to one, let us say $0.9^2$. Hence, the right-hand side would be $0.1 \times 0.05$, which is equal to $0.005$. Concerning the left-hand side, the growth accounting literature suggests that TFP growth ($\hat{A}$ in this model) is somewhere around two percent. Thus, even without any capital accumulation ($\hat{K} = 0$), the left-hand side would be $0.02$, or four times larger than the resource drag. Such rule-of-thumb calculations suggest that although the finiteness of resources imposes a drag on growth, it is very likely that output can continue to grow or at least remains constant, because technical progress and factor accumulation easily offset the resource drag.

Another important property of the Hotelling rule is that it is not only the optimal solution for the resource owner’s profit maximization problem; it is also the optimal solution for society as a whole. This should not be surprising, because the rule has been derived under optimistic neoclassical 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 be allocated

---

2 The definition of capital as a reproducible factor is a major point of disagreement between neoclassical and ecological economists, because the latter argue that 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 ongoing economic growth. See, for example, Cleveland and Ruth (1997) for more details.
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 run economic growth despite the finiteness of natural resources.

This line of reasoning has indeed been applied in the resource economics literature. Solow (1974), for example, argues that sustainable growth is possible if man-made capital is accumulated rapidly enough so that it can substitute for the declining natural capital, showing that this is possible even without technological change as long as the income share of man-made capital is larger than that of natural capital. Hartwick (1977) derives an optimal savings rule which states than an economy should invest the entire resource rent in new capital. The Hotelling rule together with the Hartwick rule has since become a cornerstone of resource economics. However, the conclusions drawn from the simple Hotelling rule rest on a number of peculiar assumptions and should therefore be treated with caution. Especially the assumption of Cobb-Douglas production is risky, because this implies that the substitution elasticity between finite resources and other production factors is equal to one. This is a knife-edge assumption; in reality this elasticity may take higher or lower values. In the 1970s a lively debate emerged between neoclassical economists, who argued that substitution between the two factors is relatively easy, and ecological economists, who argued that substitution is difficult if not impossible (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 Nonrenewable 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 basic Hotelling model 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.

Several authors have performed empirical studies to test the validity of the above-mentioned extensions of the Hotelling rule. It has to be mentioned that all such empirical tests face considerable practical hurdles, mainly because the available data are not the data that we are truly interested in. Nevertheless, the evidence against the simple Hotelling rule is overwhelming. This is why several extended versions of the Hotelling model have appeared, which we will examine in section 4.

3.1 Resource Prices

Figure 3: Real price development of four major natural resources

Figure 3 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
Should we worry about the failure of the Hotelling rule?

graphical comparability. For the most important resource, crude oil, we have information only from 1949 on. For 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 crude oil price rose by more than 60 percent, reached a plateau, and then from 1978 to 1981 it almost tripled. It then declined during the 1980s and 1990s to reach a low in 1998, when it was cheaper (in real terms) than before the first oil shock. From 1998 to 2003 it increased again.

While it would be outside the scope of this chapter to give a detailed analysis of the development of oil prices during the past few decades, and such an extensive analysis is available elsewhere, let us 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 highly 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, have caused a pronounced fall in oil prices. Second, oil prices have generally not increased 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 has been a marked increase in oil prices, which is 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. This is in gross contrast to the Hotelling rule, which predicts that the prices of nonrenewable resources should be rising at a constant rate. The prices of the other resources exhibit similar features. Over the 20th century as a whole, there has been a downward trend in resource prices. From 1900 to 2000, the prices fell by almost 40 percent in the case of zinc and copper, and by 20 percent in the case of iron ore. There have been 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

---

3 The “constant” rate of price increase stems from the assumption of a constant interest rate, which is not entirely realistic. Interest rates have generally fallen after 1982, which could explain a slower rate of price increase, but it still cannot explain falling resource prices.
Should we worry about the failure of the Hotelling rule?

resources have been falling more often than rising. This observation is by no means new. It was already observed by Prebisch (1950) and Singer (1950), who argued that the backwardness of most Third World countries was to a large extent caused by their dependence on natural resource exports, the prices of which were falling over time. To sum up, the prices of important nonrenewable resources have been volatile, especially in times of political turmoil, and they have generally fallen during the 20th century. These observations are clearly not in line with the Hotelling rule, which suggests that some of the assumptions on which the Hotelling rule is based are seriously violated in the real world.

3.2 Resource Consumption

We have shown above that on theoretical grounds an economy should consume a constant share of its remaining nonrenewable resource stock every year, and that the flow of resource consumption should be falling over time. Table 1, however, shows that this has not been the case for fossil fuels.

<table>
<thead>
<tr>
<th>World Primary Energy Demand (Mtoe)</th>
<th>1971</th>
<th>2000</th>
<th>2010</th>
<th>2030</th>
<th>Avg. annual growth 2000-2003 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coal</td>
<td>1,449</td>
<td>2,355</td>
<td>2,702</td>
<td>3,606</td>
<td>1.4</td>
</tr>
<tr>
<td>Oil</td>
<td>2,450</td>
<td>3,604</td>
<td>4,272</td>
<td>5,769</td>
<td>1.6</td>
</tr>
<tr>
<td>Gas</td>
<td>895</td>
<td>2,085</td>
<td>2,794</td>
<td>4,203</td>
<td>2.4</td>
</tr>
<tr>
<td>Nuclear</td>
<td>29</td>
<td>674</td>
<td>753</td>
<td>703</td>
<td>0.1</td>
</tr>
<tr>
<td>Hydro</td>
<td>104</td>
<td>228</td>
<td>274</td>
<td>366</td>
<td>1.6</td>
</tr>
<tr>
<td>Other renewables</td>
<td>73</td>
<td>233</td>
<td>336</td>
<td>618</td>
<td>3.3</td>
</tr>
<tr>
<td>TPES</td>
<td>4,999</td>
<td>9,179</td>
<td>11,132</td>
<td>15,267</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Source: WEO 2002

Table 1: World Primary Energy Demand in selected year (past and future projections)

The consumption of coal increased from 1,449 Mtoe in 1971 to 2,355 Mtoe in 2000. Natural gas consumption spectacularly more than doubled from 895 Mtoe in 1971 to 2,085 Mtoe in 2000, while crude oil consumption still increased by almost 50 percent from 2,450 Mtoe to 3,604 over the same time period. Thus, the consumption of these nonrenewable resources has substantially increased, not diminished. Empirical observation thus contradicts the theoretical analysis by Dasgupta and Heal. Furthermore, according to WEO estimates, consumption of
these resources is projected to increase even further between 2000 and 2030 at annual rates between 1.4 percent (in the case of coal) and 2.4 percent (in the case of gas).

Figure 4: Consumption of selected resources in the 20th century

Figure 4 shows the global consumption of three important non-renewable resources, namely iron ore, zinc, and copper. Consumption is indexed with 1949 as the base year for better comparability, and the vertical axis is plotted as a logarithmic scale. Obviously, the consumption of these resources has not fallen either. Quite surprisingly, instead of falling at a constant exponential rate, consumption of these resources appears to have grown at such a rate. There are three marked dips in resource consumption coinciding with the world wars and the Great Depression. The growth rate of resource consumption seems to have dropped in the early 1970s and remained roughly constant after that. To sum up, there is still no evidence of falling resource consumption over time, and this is in contrast with the theoretical finding by Dasgupta and Heal.

4 Reconciling the Theory with the Evidence

The previous section has shown that the simple Hotelling rule is at odds with empirical observations. The simple rule, however, was derived under some very restrictive assumptions.
Should we worry about the failure of the Hotelling rule?

In this section, we will relax a few of these assumptions and examine whether the theoretical implications become more in line with reality.

4.1 Extraction Costs

The simple Hotelling rule was derived under the assumption that resources can be extracted at zero cost. Sinclair (1994) argues that this may be an acceptable approximation in the case of oil-abundant countries like Saudi-Arabia, but in the case of oil drilling in the North Sea or Alaska extraction costs are clearly different from zero. Therefore, it may be reasonable to allow for positive marginal extraction costs. Specifically, we assume that extraction costs are proportional to output, so marginal costs \( mc \) are constant.

This extension requires only a slight reformulation of the basic Hotelling model. The Hotelling rule, as expressed in (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 costs, so (4) can be seen as a special case in which marginal cost equals zero. Thus, we can rewrite it as:

\[(4a) \quad (1 + r)(p_t - mc) = p_{t+1} - mc\]

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 is derived from Latin, meaning “in place”, so it is the value of leaving the resource in place instead of removing it. Expressed in more common words, it is the opportunity cost of extracting the resource, because extraction now means that less extraction is possible in the future. Let us simplify notation by defined \( q_t = p_t - mc \) as the in situ value of the resource in period \( t \). This allows us to write:

\[(15) \quad q_t = q_0 e^{rt}\]

We thus can 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,
of course, 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 $r$:

\[(5a) \quad p_t = q_0e^{rt} + mc\]

Since the resource price is the sum of the royalty, which is growing at rate $r$, and $mc$, which is constant, it is easy to see that the price grows more slowly than the royalty. This is shown graphically in Figure 5.

![Price and In Situ Value](image)

**Figure 5: Price and in situ value**

The price curve in Figure 5 runs parallel to the in situ curve, since price is the sum of in situ value and marginal cost. As the in situ value is growing and marginal cost are not, the ‘in situ component’ becomes an increasingly important determinant of the resource price, while the influence of the ‘cost component’ diminishes. This result will be important in the following sections.

However, although positive marginal extraction costs make the Hotelling model more realistic and add more interesting insights, they still cannot explain the falling resource prices that we observe in reality. According to the modified Hotelling rule (4a), resource prices should still be growing, albeit at a rate which is lower than the interest rate.
Should we worry about the failure of the Hotelling rule?

So far, we have assumed marginal extraction cost to be constant. This may be unrealistic. It is sometimes argued that there are stock effects, i.e. the marginal extraction cost depends on the size of the remaining stock. This can be the case in mining, for example, when the initial extraction takes place close to the surface, and over time the mines have to go deeper into the ground. The effect is that marginal extraction costs rise over time. Since the resource price is the sum of the in situ value and the marginal extraction cost, and the latter is not growing over time, prices must grow faster than under the simple Hotelling rule. Thus, stock effects in the form of rising marginal extraction cost are even more contradictory to the empirical evidence of constant or falling resource prices.

4.2 Technological Progress

In the previous section it was assumed that marginal extraction cost is constant over time, or increasing in the case of stock effects. It is also possible, however, that these costs 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 be growing at the rate of interest. The model is still similar to the above. The resource price is a combination of in situ value, which is growing over time, and marginal extraction cost, which is now falling 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 are at first falling and the rising 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. A more complicated picture arises in Farzin (1992), who focuses on the rent component of price, rather than the price itself. In addition to technological change he allows for diminishing returns to extraction and stock effects. He

4 Under diminishing returns the marginal extraction cost rises with current extraction, whereas with stock effects it rises with cumulative extraction.
finds that the time path of scarcity rent may be nonmonotonous. That is, it may be rising for some interval, then falling, then rising again, and so on. This model shows that the simple Hotelling rule of rising scarcity rents is too simple, and that U-shaped time paths may still be too simple.

If technological change is indeed the reason for the Hotelling failure, 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 will 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, which goes at the expense of other R&D efforts. TFP growth falls, and GDP growth as well. In the real world the R&D market 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 since the benefits of extraction cost reductions accrue to a small group of agents, namely resource owners, while 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.

4.3 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, our knowledge about the total available stock of resources is subject to significant uncertainty. New resource deposits are constantly being discovered by chance or through intentional exploration methods. Any new discovery increases the remaining 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
Should we worry about the failure of the Hotelling rule?

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 becomes then simply a costly activity which can be added to marginal extraction costs, and agents will base their decisions no longer 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 be increasing unless expectations are systematically incorrect.

4.4 Imperfect Competition

We have seen above that the Hotelling rule in (4) states that the present value of marginal profit in any period should be equal to that of any other period, and that (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:

\[ mr_t = p_t + p'_t x_t \]

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, which is a basic insight found in any microeconomic textbook.

For our analysis, this implies that a monopolistic resource supplier operates under a modified Hotelling rule. Substituting equation (16) for \( p_t \) in (4) yields:

\[ p_{t+1} + p'_{t+1} x_{t+1} = (1 + r)(p_t + p'_t x_t) \]
Should we worry about the failure of the Hotelling rule?

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

\[
\varepsilon_t = \frac{\partial x(p_t)}{\partial p_t} \frac{p_t}{x(p_t)}
\]  

This allows us to write:

\[
p_{t+1} = (1 + r)p_t \frac{1 + 1/\varepsilon_t}{1 + 1/\varepsilon_{t+1}}
\]

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 (4) is a special case of (18). 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 (18) 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 \( \varepsilon \) changes over time. As the price increases, we move along the demand curve to the left. Depending on the structure of the demand side, the elasticity may either increase or decrease. The growth rate of the resource price is now:

\[
\frac{p_{t+1}}{p_t} = (1 + r) \frac{1 + 1/\varepsilon_t}{1 + 1/\varepsilon_{t+1}}
\]

Depending on the development of \( \varepsilon \) over time, the resource price can increase slower or faster than under perfect competition. The monopoly will time the supply of resources such as to take advantage of the changing elasticity. When the elasticity is high, supply will be high initially, 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 Lasserre, 1988).
Since many real-world resource markets are indeed subject to imperfect competition, market structure has received considerable attention in the literature. Salant (1976) extended the analysis from a pure monopoly to the case where a dominant firm faces competition from a fringe of smaller firms. This cartel-versus-fringe model was refined by Gilbert (1978), who proposed using a von Stackelberg solution rather than a Cournot-Nash solution. The model had an undesirable weakness, namely the fact that the solution was time-inconsistent. This problem has been solved by Groot et al. (2003). The cartel-versus-fringe model has considerably increased our understanding of resource markets, but it cannot explain the failure of the Hotelling rule because it still predicts rising resource prices.
To sum up, a monopoly behaves like a competitive firm only if the demand curve is isoelastic. In all other instances it behaves different from a competitive firm, implying a market failure. Whether this raises or lowers the rate of resource extraction is theoretically ambiguous. If the demand elasticity increases over time, the monopoly extract resources too quickly, and in the opposite case it extracts too slowly. So far, there is little empirical evidence on the validity or nonvalidity of the Hotelling rule under imperfect competition (Withagen, 1998).

5 Institutional Explanations for the Failure

Traditionally most of the criticism to 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. It is one of the aims of this paper to provide two institutional explanations which may contribute to our understanding of the Hotelling failure.
5.1 Uncertain Property Rights

The simple Hotelling rule, and most of its extensions, assume 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 ‘resource curse’ (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 favor of the government, or the government itself is toppled by a revolution or a coup d’etat. Assuming well-defined property rights may therefore be too optimistic, especially in the long run.

The resource owner will form some expectation 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 $\eta$. Then, the expected present value of the resource revenue becomes:

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

Since $\eta$ 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:

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

This equation shows that with uncertain property rights the required growth rate of the resource price is lower. In fact, if $\eta$ is small enough to fulfil the condition $(1 + r)\eta < 1$ the resource price would actually be falling over time. As long as $\eta$ is smaller than one, so 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 3), but they cannot explain the increase in
Should we worry about the failure of the Hotelling rule?

consumption (Figure 4). Only in combination with some other factor, such as technological progress in exploration and extraction, they can explain the latter observation. Nevertheless, equation (4A) 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 which raises 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.

Long (1975) considers a rather similar problem, namely that of nationalization. He assumes that the resource stock is owned by a private firm, and that with a certain probability the state assumes control of that stock, possibly making a compensation payment to the firm. Long finds that since “nationalization is a private risk with no social counterpart, the extraction pattern under this kind of uncertainty tends to distort the allocation of resources”.

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 entire wars have been waged over the ownership of resources. The historical hostility between France and Germany 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 to disputes over the 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. A thorough empirical investigation may shed more light on this issue.

5.2 Strategic Interaction

Another possible deviation from the Hotelling rule may stem from strategic interaction between the suppliers and consumers of a nonrenewable 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 to develop 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 into the future. But it is C who makes the decision when to develop the technology. O can delay the arrival of the backstop technology only by influencing C’s decision. And this he can achieve by lying about the stock of remaining resources, which C has no information about.

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 period 0, O can bias C’s expectations about $R_0$. 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 to get 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 a case 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 over-estimated. 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 not necessarily a 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 failure 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. If new discoveries occur, people will form 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 consumed too high, and under pessimistic expectations it would be too low. In theory, therefore, the bias can go towards either direction.

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. If a successful coup is likely to occur in the near future, he has an incentive to exploit the resource as fast as possible, sell it on the world market, and
store the revenue at some safe place in a foreign country. 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 optimal.

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 be rising due to stock effects or falling 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 for 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 the property rights suggests that the government itself is either unwilling or unable to solve the problem, because otherwise they would have established these property rights already. And with strategic interaction between consumers and suppliers, a solution is also difficult to find. Since most nonrenewable 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 easily at hand. Nevertheless, it is important to acknowledge the presence of these market failures. They suggest that nonrenewable resource consumption indeed occurs too quickly, and other policy measures which reduce the resource extraction rate may therefore be welfare-enhancing.
Note that in this paper no mention has been made of the negative environmental externalities of fossil fuel consumption. Even so, we have seen that the market for these resources is characterized by a number of market failures, and that resource extraction probably occurs too quickly. If we take the external effects of resource consumption into account, especially those that lead to stock externalities, the welfare effect of these market failures is likely to be even more pronounced. The existence of stock externalities, such as CO2 accumulations in the atmosphere, would require a smaller resource consumption rate to ensure a social optimum, and a market which does not take these externalities into account will result in an extraction rate which is too high from a social welfare perspective.

Concerning the relationship between pollution and growth, the existence of nonrenewable resources does not weaken the link. If the Hotelling rule held, resource consumption would be falling 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.
8 References


Should we worry about the failure of the Hotelling rule?


