# Capacity Constraints, Expansion Options and the Optimal Extraction of Exhaustible Resources

**Hamed Ghoddusi***
**November 2008**

**Incomplete, Comments Welcome**

**Abstract:** This paper extends the literature of exhaustible resource economics by examining the investment decision of an active exhaustible resource monopolist. With demand uncertainty and endogenous price dynamics, the monopolist chooses both the production rates and the time to build an extra capacity optimally. The capacity expansion option for such a firm is modeled as a two-dimensional option on demand shocks and remaining reserves. Using a discrete-time simulation, first the dynamics of option prices and its sensitivity to different parameters is calculated. Furthermore, it is shown that the consideration of option value in investment decisions will lead the producer to choose a more conservative expansion policy and therefore causes higher prices in strong demand shock periods. Finally, the optimal production rate of the producer will change by the introduction of option feature to the problem. The findings of this paper may explain why we do not observe in practice the predictions of Hotelling rule regarding increasing prices and decreasing production rates.

**JEL Classification: Q31, G13**
**Keywords:** Capacity Expansion Options, Natural Resource Economics, Real Options

---

* Vienna Graduate School of Finance (VGSF), Heiligenstadterstrasse 46, 1190 Vienna, Austria. Phone: +43 (1) 31336 6321, e-mail: Hamed.ghoddusi@vgsf.ac.at, web: www.ghoddusi.com

I very much appreciate my supervisor Prof Engelbert Dockner for his guidance when I was writing this paper. Without his ideas I would have not been able to formulate this research in a proper way. I am also very grateful to Thomas Dangl, Damir Filipovic , Kristian Miltersen, Klaus Ritzberger, Eduardo Schwarz, Sheridan Titman, Franz Wirl and the participants of VGSF maxi conference for their kind help and valuable comments. The generous financial support by the Vienna Graduate School of Finance is sincerely acknowledged. The usual disclaimers apply.

# 1. Introduction

Since the introduction of Hotelling's (1931) formulation for the optimal extraction path of exhaustible resources, several variations and refinements of this problem have been proposed and discussed in the literature. Consideration of variable production costs, various uncertainties, oligopolistic competition, capacity limits, backstop technology, deposits multiplicity and the investment in exploration are some major examples of the these variations.

To apply Hotelling's formulation into the real world problems one needs to include different uncertainties (regarding for instance to the demand process, production and investment costs, time to build, geological properties and deposits, unexpected stops, etc) on the one hand and the capacity constraints on the other hand. It is also important to consider these two facts simultaneously and interrelated to each other. This motivates a natural stage to apply the concept of capacity expansion options in order to analysis the investment and production decisions of the producer.

As a real case, consider a monopolist producer of a natural resource (e.g. oil or copper), which faces stochastic demand and is restricted by a rigid capacity limit while possessing a capacity expansion option at the same time. The capacity limit is determined by several factors such as the number of wells, the size of tunnels, the capacity of material handing systems and the number of trained human resources. Since the production rate is restricted by these factors, the optimal extraction path of this producer will be different from what the basic Hotelling rule - or *r-percent* rule - suggests. This is simply because an optimal extraction programme in some periods may require extraction rates which are higher than the practically feasible level. This immediately inserts a further constraint to the optimal control problem of the producer.

On the other hand, the existence of a capacity expansion option provides the chance to ease this limit to some extent by bearing particular investment costs (both in the beginning to *buy* the option and in future to exercise it). These investment expenses should be treated as irreversible fixed costs. This is a plausible assumption as investments in mining and oil industry (for example the cost to drill an oil well or to build a platform or to construct tunnels in mining) are highly tailor-made and specific for the particular oil/gas field or mine. Moreover, even if in the case of some general purpose machinery and equipments, disentangling, disassembling and moving them to another site or plant might be too costly. Therefore it is acceptable to treat them as fully irreversible costs.

Because of this very irreversibility, the option to wait has a positive value and the capacity expansion option will be exercised only when the immediate exercise is profitable enough meaning that the investment benefits justifies the scarification of the option to wait. As an example, look at the case of a large producer of oil that has the right to use an oil field. The problem for this producer will be the optimal time (or to be technically more precise, the level of certain state variables) to start drilling a new well in order to add a specific amount to the current production capacity.

Although the option value for an inactive producer has been examined in the literature extensively, only one study – to the best of my knowledge - has studied the production and capacity building decisions simultaneously. Previous studies (Brennan and Schwartz (1985), Morck et al (1989), for instance) usually calculate the real option value of investing in resources with an exogenously specified *price* process. Moreover, they usually assume a right to exercise an investment option for a currently *inactive* natural resource producer. My model differs from them through the important fact that the producer is active in both periods before and after exercising the capacity option. Therefore, there is a second dynamics in the model which is the level of remaining reserves. In this sense the capacity option problem considered is a two-dimensional one.

The closest paper to the current one is Carlson, Kokher and Titman (2007) which introduces adjustment costs to the resource extraction problem in a competitive market and leaves the oligopoly case as an open question. In addition, their paper does not assume a rigid capacity constraint and only impose adjustment costs. To address the case of oligopoly, one needs to first develop a theory for a monopoly situation and this is what the current paper does. It should be noted that the current paper is one early outcome of a larger research project which aims to explore the interaction of real options and natural resource economics in a more realistic oligopolistic setting.

The value of a capacity expansion option for a monopolist (or close to monopolist) producer of an exhaustible resource such as oil or copper is an important parameter in making the capital budgeting and investment decisions. Making use of this option in the future requires some basic infrastructure to be built in the early stages of construction. For instance consider a copper producer which designs and implements facilities to extract 500,000 tons of copper per year. The producer, however, foresees that it might be optimal in future – because of stronger demand or lower production costs – to add another 500,000 units to the production by installing some new capacities to flotation or refinery processes.

This is possible only if the structure of the ore extraction system or the tunnels were built in such a way from the onset to support a 1,000,000 ton production line. Since the structures for 1,000,000 and 500,000 tons/year differ in cost, the decision to choose one of them depends crucially on the value of the (future) capacity expansion option. One can see this extra cost of building a stronger structure, which may seem for the time being useless and sunk, very similar to the money paid to buy a call option on a volatile asset. As we know, the equilibrium price of the financial option will not exceed its expected pay-offs (under risk-neutral measure), the same logic applies here for the investor in natural resources.

Moreover, both from a theoretical and practical point of view it is interesting to know about the reaction of the producer to the existence of this capacity expansion option. The producer – bounded by limited capacity – will choose an optimal rate of extraction according to a solution of a stochastic optimal control problem. This is an old observation going back to the 1980s. I extend the problem by adding a capacity expansion option as a further feature and then ask what happens if the producer is forward looking and takes into account the possibility of adding to the extraction capacity in future in her current production decision.

The previous paragraphs illustrate three major questions addressed in this paper which are:

1) What is the value of a capacity option for an exhaustible resource producer and when will it be exercised?

2) What is the impact of a capacity option on the dynamics of price?

3) Does the presence of a capacity option change the production plan of a producer? In other words, does it make any difference in supply side responses to demand shocks?

The contribution of the current paper is twofold. I first make the price dynamics endogenous by looking into the responses of the supply side to demand shocks and the change in price after exercising the option. The second contribution is more methodological. A real option problem with two state variables where the second one is implicitly dependent of the first one through another optimal control problem has been solved.

The insights from this formulation can be applied to for instance to optimize capacity expansion decisions in an environment where the underlying demand for the product is deteriorating. One example would be the case of capacity expansion for a producer introducing a new durable product whose market will shrink over time and the rate of drop in demand is a function of the rate of supply in the previous periods. Unlike other capacity expansion papers (for example Dixit and Pindyck (1997), Dangl (1999)) that assume an unlimited access to the factor market and impose the production capacity as the only limitation, in this paper the total supply of main production factor over a (not necessarily finite) time horizon is given and fixed.

The paper gives intuition over the behavior the option price when the instantaneous production rate is both dependent and independent of the realization of demand shocks. It will be shown that if we assume no production costs then with iso-elastic demand the optimal extraction rate is independent of demand shocks while production with linear demand will depend on demand shock. I, intentionally, assume zero costs in the case of isoelastic demand in order to present both classes of the problem. The most important difference will be that when the optimal production is not being affected by demand shocks, the option will be valuable only for a finite time and its value drops to zero as soon as the level of stock reaches the critical level. In contrast, with linear structure the option may live longer (still not until infinity) because of the possibility of a very high demand realization and a strong supply response.

## *2. The Review of the Literature*

With no doubt, Hotelling's (1931) paper is the most seminal work in the whole literature of natural resource economics. This paper raised a critical question that "what is the optimal rate of extracting for a particular exhaustible resource". The paper answered this question by introducing the famous r-percent rule implying that price net marginal cost (marginal revenue) will increase with interest rate for competitive (monopolistic respectively) market. As an immediate implication of this rule we should observe the futures contracts of commodities always in contango. Furthermore, the production rate should diminish over time if the demand is isoelastic. Although Hotelling refers to the issue of limited production capacity, he does not explicitly incorporate it in his model. Therefore the results are valid only for an unlimited rate of extraction. Moreover, Hotelling does not consider any uncertainty in the problem whereas the reality includes various uncertainties as explained before.

Pyndick (1980) was the first to formulate a version of Hotelling's problem with continuous stochastic demand process. Assuming an Ito process for both demand and resource depletion dynamics, he reached a HJB equation which characterizes the optimal value function and extraction trajectory. He shows that if the marginal cost does not depend on reserves, the optimal extraction rate of the stochastic problem will be the same as the deterministic one. In other words, the *r-percent* rule for competitive and monopolistic case holds even for this case under expectations operator.

Cambpell (1980) introduced capital intensity based capacity constraint into the model. By this further assumption, the production at each instant is limited to a maximum level which is a function of net accumulated capital. Solving the constrained dynamic optimization problem he showed that the capacity constraint will be binding at some initial periods and then will be relaxed. Therefore it is optimal to build the capacity only at the very beginning of production. This comes from the fact that the optimal production trajectory is decreasing is time (or resource level), hence the shadow price of extra capacity is also decreasing over time and becomes zero after some time.

Davis and Moore (1998) revise the Hotelling valuation rule based on capacity constraints. The older rule suggests that for in a competitive market the present value of a natural resource with known quantity is just the product of net value (price less production costs) and the amount of resource. Nevertheless, the empirical tests concluded that the value in reality is lower than the predictions of this rule. Davis and Moore justify this difference by introducing capacity constraint and show that as long as the capacity is binding, the value function is lower than the unlimited case. They however do not provide an explicit valuation formula for the constraint case.

Litzenberger and Rabinowitz (1995) use the concept of (real) options existing in oil production to explain the seemingly odd backwardation in oil futures contracts. Backwardation is not consistent with the predictions of Hotelling rule as it expects the future prices to rise and therefore futures curve to be in contango. They argue that because of the value of this real option any decision to produce right now means to scarify the option to wait. As a result current production will happen only if the present value of future incomes plus the option value exceeds the value of spot production. Since the option value is always positive, the present value of future income should be lower than the spot value which is equivalent to strong backwardation.

## *3. The Model*

The risk-neutral monopolist is endowed with an initial level R0 of the deposit of exhaustible resources. I assume no uncertainty regarding the resources and no exploration effort, therefore the deterministic dynamics of resource depletion is given by $dR_t = -q_t dt$ where $q_t$ is the instantaneous rate of production and $R_t$ the level of remaining reserves. An equivalent integral formulation would be $R_0 = \int_0^\infty q_t dt$

The inverse demand is a function of production rate $\boldsymbol{q_t}$ and an exogenously given random process $\boldsymbol{X}$ and is written as $P_t = f(X, q_t)$. That extraction costs are neglected. This is a realistic assumption for example for the production of oil at the Middle East where the variable cost is around 2 US$ per barrel. Therefore the profit is determined by $\pi(q_t) = p(X, q_t).q_t$. The stochastic demand parameter X is a geometric Brownian motion with the following dynamics: $dX = X.\mu.dt + X.\sigma.dW$

The producer maximizes the sum of the present value of future profit streams which is equivalent to the next continuous-time stochastic optimal control problem:

$$\begin{aligned}
&V(R_0, X_0) = \underset{q}{Max}\, \mathrm{E} \int_0^\infty \pi(q_t) e^{-rt} dt \\
&s.t. \\
&\dot{R} = -q \\
&dX = X.\mu.dt + X.\sigma.dW \\
&R_t, q_t \geq 0
\end{aligned}$$

In this problem the interest rate is taken constant and a *dot* on the variables shows the time derivative.

### 3.1) Inverse Demand Function and the Impact of Shocks

To analysis the behavior of the exhaustible resource monopolist, one can specify different functional forms for the inverse demand and the profit function and the optimal production trajectory will change accordingly. Two widely used functional forms in this area are the isoelastic demand function specified as $P(q, X) = q^\gamma.X, \gamma < 0$ and the linear demand function given by $P(q, X) = X - q, X > 0$ where X represents an exogenous demand parameter (usually interpreted as the price of backstop technology) and q accounts for the production rate. It can be shown and will be discussed later that with isoelastic demand the optimal extraction rate is **independent** of the realization of demand shocks while with the linear demand the shocks change the optimal instantaneous extraction rate. Let us motivate this fact by a simple example.

**Example 1:** Consider an exhaustible resource monopolist which faces a two-period problem. The initial reserve of resource R is given (R=100) and the interest rate is 10%.

The risk-neutral producer solves the following problem

$$\begin{aligned}&\underset{q1,q2}{Max}\,\pi(q_1)+e^{-0.1}\pi(q_2)\\&st\\&q_1+q_2\le 100\\&q_1,q_1\ge 0\end{aligned}$$

Solving the problem one sees that for the interior solution the optimality condition is $MR_2 = e^{0.1}MR_1$. Now look at this equation with two different forms of inverse demands:

First by isoelsatic demand function:
$MR(q_2)=(1+\gamma)Xq_2{}^{\gamma}=e^{0.1}(1+\gamma)Xq_1^{\gamma}\Rightarrow q_2=q_1(e^{0.1})^{1/\gamma}$, Independent of X!

And now plug the linear demand function:
$MR(q_2)=(X-2q_2)=e^{0.1}(X-2q_1)\Rightarrow q_2=-0.5*r*X+e^{0.1}q_1$ , Function of X!

Since each demand function represents certain properties in the real world, both forms will be studied in this paper.

### 3.2) Optimal Extraction Rates

It is beneficial to first illustrate different optimal policies under different assumptions concerning the behavior of the demand. Both forms of the demand functions with deterministic and stochastic dynamics are presented in the next sub-sections.

#### 3.2.1) Deterministically Growing Iso-elastic Demand

The simplest case would be the one where the demand is given by $P(X,q)=X.q^{\gamma},0<\gamma$ and the cofficient X has a deterministic dynamics as $dX = X\mu dt$

Basic Hotelling's rule suggests that in the absence of production costs the marginal revenue of the monopolist will grow with the interest rate. Using this rule:

$$\frac{dMR}{MR}=rdt\ ,\ \mathrm{MR}\ =(1+\gamma)\mathrm{X.\ q}^{\gamma}\ ,\ \mathrm{dMR}\ =(1+\gamma)[\mathrm{dXq}^{\ \gamma}+X\mathrm{d}(\mathrm{q}^{\ \gamma})]=$$
$$(1+\gamma)[\,q^{\gamma}X\mu dt\ +\ X\gamma q^{\gamma-1}dq\,]\Rightarrow$$

$$\frac{dMR}{MR}=\frac{(1+\gamma)[\,\mathrm{q}^{\gamma}X\mu dt\ +\ X\gamma q^{\gamma-1}dq\,]}{(1+\gamma)\mathrm{X.\ q}^{\gamma}}=rdt\ \Rightarrow\ \gamma\frac{dq}{dt}=(r-\mu)q$$

$$\Rightarrow\ \frac{\dot{q}}{q}=\frac{r-\mu}{\gamma}\Rightarrow\ q_t=q_0e^{\frac{r-\mu}{\gamma}t}$$

$$\int_0^{\infty}q_0e^{\frac{r-\mu}{\gamma}t}=R_0\Rightarrow\ q_t=R_t(\frac{r-\mu}{-\gamma})$$

As seen the production rate is always a linear function of the remaining reserves and declines exponentially.

**3.2.2) Stochastic Iso-elastic Demand**

Pindyck (1980) shows that even with stochastic demand shocks the basic Hotelling's rule for marginal revenue (price in the case of competitive market) holds. Therefore we have:

$$\frac{EdMR}{MR} = rdt \text{ , MR} = (1+\gamma)\text{X.q}^{\gamma} \text{ , EdMR} = (1+\gamma)\text{E}[\text{dXq}^{\gamma} + X\text{d}(\text{q}^{\gamma})] = (1+\gamma)[q^{\gamma} X\mu dt + X\gamma q^{\gamma-1} dq] \Rightarrow$$

$$\frac{EdMR}{MR} = \frac{(1+\gamma)[\text{q}^{\gamma} X\mu dt + X\gamma q^{\gamma-1} dq]}{(1+\gamma)\text{X.q}^{\gamma}} = rdt \Rightarrow \frac{\dot{q}}{q} = \frac{r-\mu}{\gamma} \Rightarrow q_t = q_0 e^{\frac{r-\mu}{\gamma}t}$$

$$\int_0^{\infty} q_0 e^{\frac{r-\mu}{\gamma}t} = R_0 \Rightarrow q_t = R_t (\frac{r-\mu}{-\gamma})$$

This is important to notice that the instantaneous production rate does not depend on X (meaning that it is independent of the realizations of the demand shocks) and is only the function of parameters of demand function.

**3.2.3) Deterministic Growing Linear Demand**

Unlike the isoelastic case the optimal production with linear demand will not be independent of demand shocks.

$$MR(t) = MR(0)e^{rt} \Rightarrow (X(t) - 2q(t)) = (X(0) - 2q(0)).e^{rt} \Rightarrow$$

$$q(t) = \frac{1}{2}[X(0)(e^{\mu t} - e^{rt})] + q(0)e^{rt}$$

With linear demand the reserves will deplete in finite time (if the growth rate of demand is less than the interest rate) therefore there is a full depletion time T where the production stops there and extraction rate is zero for any t>T. On the other hand the total supply of reserves is given by R and the value of shock at time 0, X(0) is known. Using these facts one can write down two equations with two unknowns (*q(0)* and *T*) in order to fully characterize the production path.

$$\int_0^T q(t) = \int_0^T \frac{1}{2}[X(0)(e^{\mu t} - e^{rt})] + q(0)e^{rt} = R$$

$$\text{q(T)} = \frac{1}{2}[X(0)(e^{\mu \text{T}} - e^{r\text{T}})] + q(0)e^{r\text{T}} = 0$$

The production rate for the linear demand may tend to grow over time and therefore the optimal production policy can be downward or upward sloping (depending on the ratio of interest rate and the growth rate of demand). If the growth rate of demand is lower than the interest rate the reserves will be depleted in finite time while with isoelastic demand it will never be fully depleted. In the next sub-section the samples of the optimal production patterns together with the optimal rates under capacity constraints will be demonstrated.

### 3.3) The Impact of Capacity Constraints

As explained before, in reality the extraction rate of natural resources is limited by installed and active capacity. As it is extremely capital intensive the production capacity is usually fixed for a period of time. For instance, in the oil industry building new capacity may take between 3 to 10 years. The problem of producer therefore is a constrained optimization problem.

If the demand is isoelastic the extraction rate in a deterministic environment is declining over time. Based on Campbell (1980) the capacity constraint is binding only for some initial period of production horizon and once the resource reaches a critical level it will stop binding forever. This helps to write down the value function of resource production problem with binding capacity as the sum of two separate sub-problems: the problem for the period when the constraint is binding and the problem for the period right after that. In the first period there will be a fixed rate of production. Therefore the value of first period is just the present value of a fixed income stream in a finite horizon. The second period problem starts at T* and its current value can be calculated by taking the sum of discounted future cash flows coming from the optimal control problem. An example of extraction trajectory with and without capacity constraint is depicted in the figure 1.

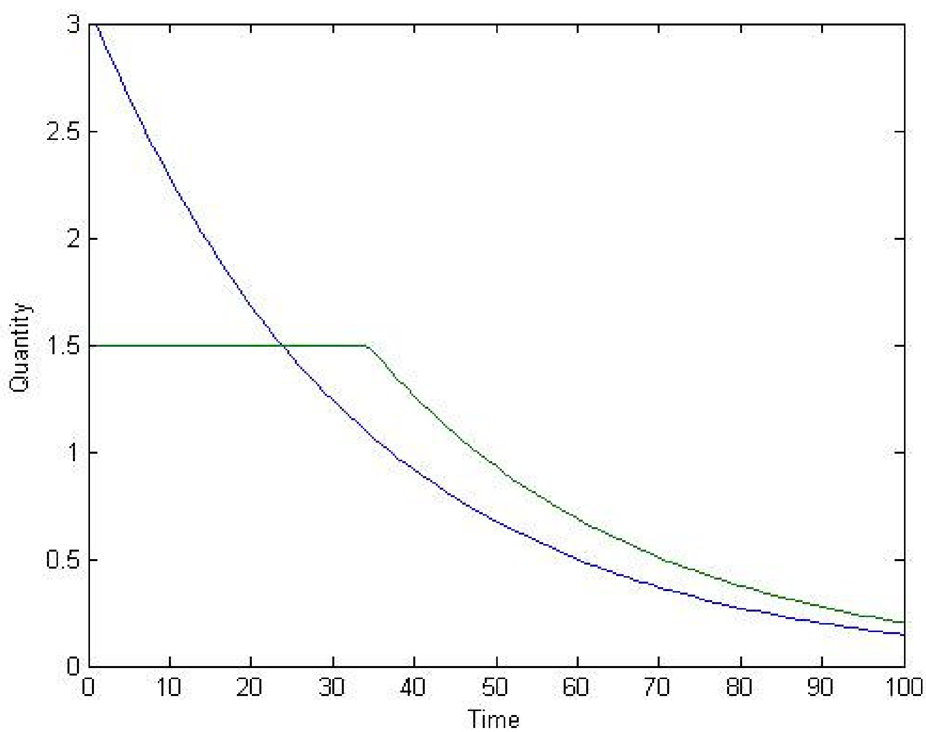



**Figure 1: Optimal trajectory of q for unconstrained (blue) and constrained (green) case and iso-elastic demand**

The graph gives a graphical intuition of the previously discussed behavior. The optimal extraction rate is fixed for the constraint problem and once the level of reserves reaches a certain value the dynamics of extraction rate is similar to the unconstraint problem.

On the other hand, if the inverse demand has a linear form, then the dynamics of production would take different shapes depending on the drift of demand process. If the demand grows slower than the interest rate then the production starts from the beginning and declines until the resource is fully depleted (this is true only if there is no production costs otherwise the producer may leave some resources underground forever). Adding capacity constraint just makes the time of depletion longer. On the other hand, if the demand grows faster than the interest rate, then it is optimal to wait and benefit from a stronger demand in future periods. If the time horizon is infinite then the resource will never be extracted and this leads into a bubble in the value function. With a finite time horizon the dynamics of production would look like to the right panel of figure 2 where the producer starts extraction in finite time but with an upward extraction path.

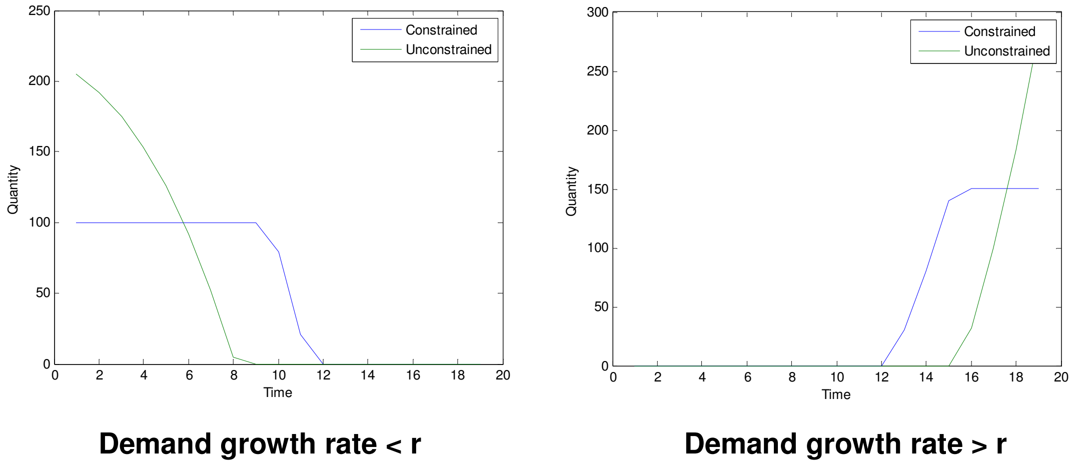



**Figure 2: Optimal trajectory of q for unconstrained (blue) and constrained (green) case and linear demand**

In the next step, the stochasticity is introduced to the demand process. As discussed before, under linear demand the instantaneous rate will depend on the realization of the shocks. Figure 3 first shows a sample of production rates under unconstraint problem and two different levels of capacity constraints. It can be inferred from the graphs that when the capacity constraint is too low the time-path of optimal production (in finite time) will be flat. This is a direct result of forward looking behavior. The producer knows that the production will terminate in a finite period and the total production during the whole future (starting from the next period) will not exceed a certain amount. This destroys the inter-temporal optimization nature of the problem because the shadow value of more units of reserves in these periods is zero. As a result, it is optimal to produce according to the results of a static optimization problem which in this case turns out to be equal to the maximum capacity. This result may justify to some extent why we do not observe the declining pattern of production in reality.

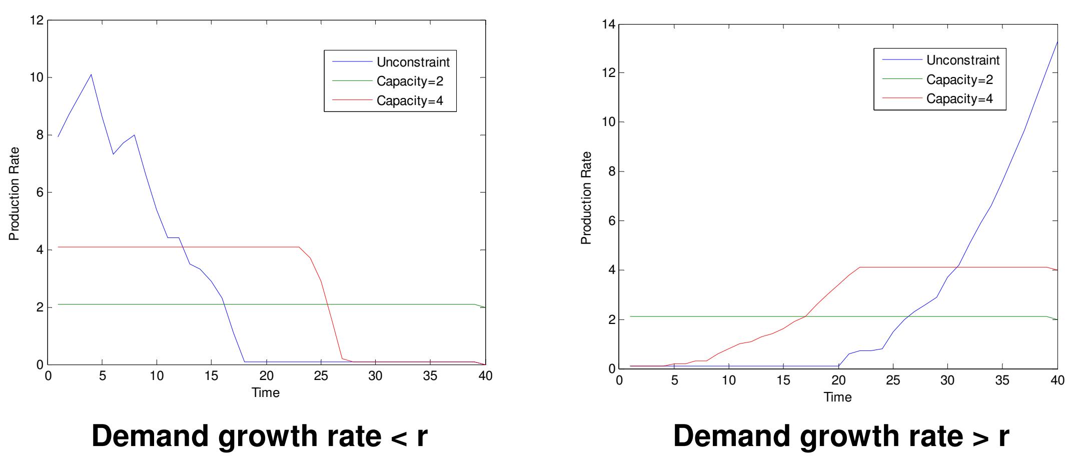



**Figure 3: Optimal trajectory of q for stochastic demand**

### 3.4) The Value of Extra Unit of Production for Iso-elastic case

A capacity expansion option for an active producer is similar to a one-time American switching option where the holder can shift between two securities with a given cost *I*. As mentioned before, the option considered here has two unique features which make it different from other capacity options. First, for a producer of exhaustible resources the total supply of resources is given and fixed. This means if the producer postpones the investment in the capacity option for one period, the profits of that period will not be lost completely but will contribute to the firm value in the next period. This is different from the way the usual capacity option problems are formulated. The typical way to motivate a capacity expansion option is to introduce a *widget* producer which takes prices as given and determines when to invest in further capacity. If the capacity building decision is delayed for one period, the operational profit of that period is gone and therefore it is costly to postpone investment decision. Because of this cost the producer tends to exercise its option in a finite time. This is not the case for the problem of the current paper since natural resource monopolist considers the inter-temporal characteristic of the problem.

The second major difference relies on the fact that the reserves available for the producer are being constantly depleted. Therefore, at any time period in future the level of reserves will be lower than today. Ceteris Paribus, it means that the *current* value of a capacity expansion option will be lower than today's current value at any future period because the second underlying has declined. This will be discussed in detail in the next sections.

To approach the valuation of capacity option one notices that introducing a new constraint to an optimization problem will never improve the value function. Moreover, if the solution to the optimal control problem is unique and the constraint is binding over an interval with non-zero measure, the value function will be lower compared to the unconstraint problem. The problems considered in this paper all have concave objective functions and convex constraints spaces. Therefore, their solution is unique and as a result the value function is strictly decreasing in the difference between maximal capacity needed and given capacity limit.

Figure 4 depicts the sensitivity of value function to capacity constraints for a given set of parameters. It can be inferred from the graph that relaxing the capacity constraint up to certain point will improve the value function and after that point the value of constraint and unconstraint problems will be the same.

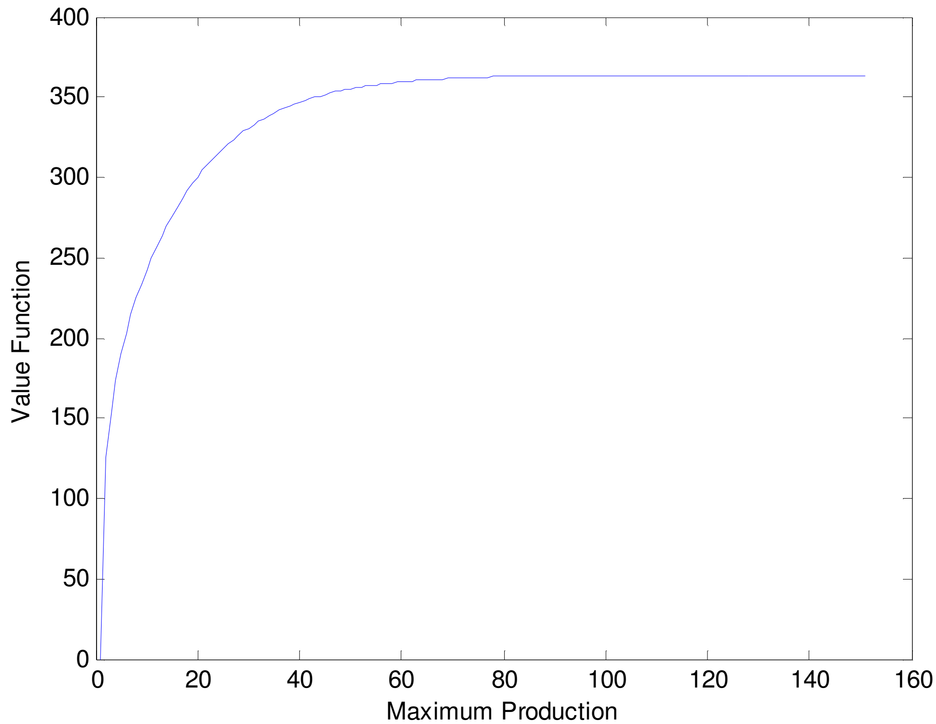



**Figure 4: Impact of Capacity Constraints on the Value Function**

It is easy to find this *"non-binding"* point for the isoelastic demand function. The equation in the sub-section 4.2 gives the relation between the remaining available resources and the instantaneous optimal production. Using this relation one can characterize the optimal path for the case of limited capacity. The ideas is to divide the extraction path into two regions where in the first region the constraint is biding and therefore the flat production rate will be equal to the capacity. In the second region where the capacity is not binding and will not bind anymore the problem is the usual Hotelling problem without capacity constraints.

For a given production capacity $\overline{Q}$ the capacity will not bind if and only if:

$$q_t = R_t(\frac{r-\mu}{-\gamma}) \leq \overline{Q} \Rightarrow R_t \leq \overline{Q}(\frac{-\gamma}{r-\mu})$$

If the initial reserves exceed that threshold, the production curve will be flat and equal to $\overline{Q}$ at the beginning. Using this fact one can calculate the time it takes for a producer with a given reserve $R_0$ to reach the non-binding region: $R_0 - t^*\overline{Q} = \overline{Q}(\frac{-\gamma}{r-\mu}) \Rightarrow t^* = \frac{R_0}{\overline{Q}} + \frac{\gamma}{r-\mu}$

Based on these calculations the value function of the producer is given by:

$$V(R_0,X_0,\overline{Q}) = V_{binding} + e^{-rt^*} V_{unconstrained}$$

$$V_{binding} = E\int_0^{t^*} \pi(\overline{Q})e^{-rt}dt = \int_0^{t^*} E(X_t)\overline{Q}^{\gamma+1}e^{-rt}dt = \int_0^{t^*} X_0 e^{\mu t}\overline{Q}^{\gamma+1}e^{-rt}dt = X_0\overline{Q}^{\gamma+1}\int_0^{t^*} e^{(\mu-r)t}dt = \frac{X_0\overline{Q}^{\gamma+1}}{\mu-r}[e^{(\mu-r)t^*}-1]$$

$$V_{non-binding} = E\int_0^{\infty} \pi(q_t)e^{-rt}dt$$

$$\pi(q_t) = X'_t(q_t)^{1+\gamma} = X'_0 e^{\mu t}(q_0 e^{\frac{r-\mu}{\gamma}t})^{1+\gamma} = X'_0 q^{1+\gamma}{}_0 e^{rt+\frac{r-\mu}{\gamma}t} \Rightarrow e^{-rt}\pi(q_t) = X'_0 q^{1+\gamma}{}_0 e^{\frac{r-\mu}{\gamma}t} \Rightarrow E\int_0^{\infty}\pi(q_t)e^{-rt}dt = X'_0 q_0^{1+\gamma}(\frac{-\gamma}{r-\mu})$$

At the begining of non-binding region: $q_0 = \overline{Q}$ and $X'_0 = X_0.e^{\mu t^*}$

By plugging the parameters back to the valuation equation one gets:

$$V(R_0, X_0, \overline{Q}) = \frac{X_0\overline{Q}^{\gamma+1}}{\mu-r}[e^{(\mu-r)t^*}-1] + e^{-rt^*}(\overline{Q})^{1+\lambda}X_0(\frac{-\gamma}{r-\mu}) = [\frac{X_0\overline{Q}^{1+\gamma}}{\mu-r}]((1+\gamma)e^{(\mu-r)t^*}-1)$$

To find the value of an extra unit of capacity *VC* denote the critical level of reserves where the capacity constraint just stops binding by R*. Use the fact that under the isoelastic demand, production rate is always decreasing therefore we can specify two regions:

$$1) R \le R^* \Rightarrow VC = 0$$
$$2) R > R^* \Rightarrow VC = V(R, X, \overline{q}_1 + 1) - V(R, X, \overline{q}_1)$$

The first equation simply says that when the reserve reaches the non-binding threshold the capacity constraint will never bind. Therefore an extra unit of capacity has no value. The second equation tells that at any point before this region, one unit of extra capacity will relax the capacity constraint and improves the value function. The difference between the values of two problems is the value of an extra unit of capacity which is equal to the shadow value of capacity constraint.

Figure 5 shows the time path of the value of one unit of extra capacity for different assumptions about demand process. The dot line refers to the case of constant demand and therefore the value of extra unit of capacity is constantly declining over time. This conforms to the results of Cambpell (1980) suggesting that the optimal time to invest in extra capacity is the beginning of the production.

The dashed and continuous lines represent the case of determinist growing demand and a demand with GBM process and positive drift. It is an important observation to notice that although the change in demand parameter *X* does not influence the instantaneous production rate, it still has a strong effect on the total value of the firm through changing the profits.

Therefore one sees that the change of demand influences the value of an extra unit of capacity significantly.

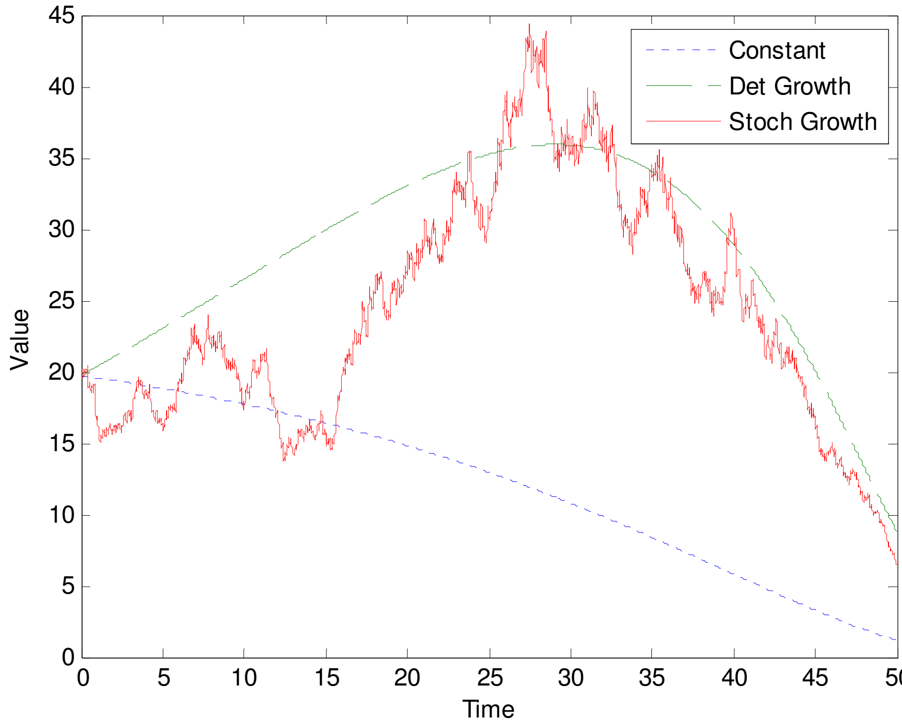



**Figure 5: Time Path of an Extra Unit of Capacity**

As seen in figure 5 the value of extra unit of capacity converges to zero after certain time. This comes from the fact that the reserves are constantly depleting over time and when the level of reserves reaches the critical threshold where the capacity constraint is not binding, extra capacity has zero value. This result may change if one introduces uncertainties about the level of reserves through adding for instance a jump process accounting for the discovery of new reserves. Under this new setting there is always a positive probability that the reserves becomes larges in future and therefore the value of extra unit of capacity will not die and remains always positive. I abstract from this further aspect and leave it to the extensions of the paper.

### 3.5) The Firm's Problem under Capacity Constraints

A producer with arbitrary demand function and a given non-depreciating capacity constraint solves the following stochastic optimal control problem where X and R are state variables, q is the control variable and *qmax* refers to the maximum extraction rate:

$$V = \max_q E\int_0^\infty \pi(q_t).e^{-rt}dt = \max_q E\int_0^\infty q.p(q).e^{-rt}dt$$

$$s.t.$$

$$\dot{R} = -q$$

$$0 \le q_t \le q_{\max}$$

$$dX = X.\mu.dt + X.\sigma.dW$$

$$R \ge 0$$

Using standard arguments one gets to the following Hamilton-Jacobi-Bellman equation which characterizes the value function:

$$rV = \underset{q \le q\max}{Max}\{\pi(q) - qV_R + X\mu V_x + \frac{1}{2}\sigma^2 X^2 V_{XX}\}$$

$$V(0, X) = 0$$

$$V(R,0) = 0$$

If the production rate depends on X, one can distinguish two cases for q. The first case associates with all realizations of X where $q_t \ge q\max$. In this case the maximum occurs at *qmax* boundary and the corner F.O.C plugged backed into HJB equation leads to the following PDE

$$rV - \pi(q\max) - q\max V_R + X\mu V_x + \frac{1}{2}\sigma^2 X^2 V_{XX} = 0$$

$$V(0, X) = 0$$

$$V(R,0) = 0$$

One should notice that at any point where capacity constraint is not binding the option will not be exercised. This comes from the fact that the shadow value of extra capacity is zero at the point and therefore adding extra capacity will not improve value function while will impose an investment cost of *I* and as a result is a sub-optimal decision.

On the other hand if the capacity constraint is not binding FOC condition gives the interior solution which in the case of linear demand will be:

$$\frac{\delta}{\delta q}\pi(q) - V_R = 0 :\Rightarrow X - 2q^* = V_R \Rightarrow q^* = (X - V_R)/2$$

Substituting it back into HJB equation gives value PDE and its boundary condition:

$$rV - 0.5(V_R - X)(X - V_R) - 0.5(X - V_R)V_R + X\mu V_x + \frac{1}{2}\sigma^2 X^2 V_{XX} = 0$$
$$V(0, X) = 0$$
$$V(R,0) = 0$$

These PDEs are not easy to solve analytically. Therefore we have to use numerical methods to evaluate value function for any given set of (R,X).

**3.6) The Value of Capacity Expansion Option for Iso-elastic Demand**

Now consider the problem of the resource producer with capacity expansion option at her disposal. As discussed before, this option is like a perpetual American option with exercise price of I. The important point is to notice that in our case this is a two-dimensional option whose value is driven by two state variables *X* and *R*. The key question is to find the free boundary as an implicit function of $J(X^*, R^*) = 0$ or a curve on the state space (X,R) which characterize the region where exercising the option is optimal.

To calculate the value of capacity expansion option we consider the basic no arbitrage equation. As explained before we do not need to consider dividend effect as the value function captures this effect:

$$\rho F dt = \varepsilon(dF) = (F_x.\mu.X + \frac{1}{2}\delta^2 X^2 F_{xx} - q^* F_R)dt \Rightarrow F_x.\mu.X + \frac{1}{2}\delta^2 X^2 F_{xx} - q^* F_R - \rho F = 0$$

The initial and boundary conditions of PDE will be:

$$F(R,0,q) = 0$$
$$F(0, X, q) = 0$$
$$F(R, X^*, q_1) = V(R, X^*, q_2) - V(R, X^*, q_1) - I$$
$$F_x(R, X^*, q_1) = V_x(R, X^*, q_2) - V_x(R, X^*, q_1)$$
$$F_R(R, X^*, q_1) = V_R(R, X^*, q_2) - V_R(R, X^*, q_1)$$

The first condition comes from the fact that X=0 is an absorbing state for GBM process. The second condition suggests that when there is no reverses the option has no value. The rest are the usual value matching and smooth pasting conditions for two variables case.

Unfortunately this PDE is not feasible to solve rigorously. Therefore I use binomial tree method to compute the value of the option. Figure 6 shows the change of option value versus changing different parameters.

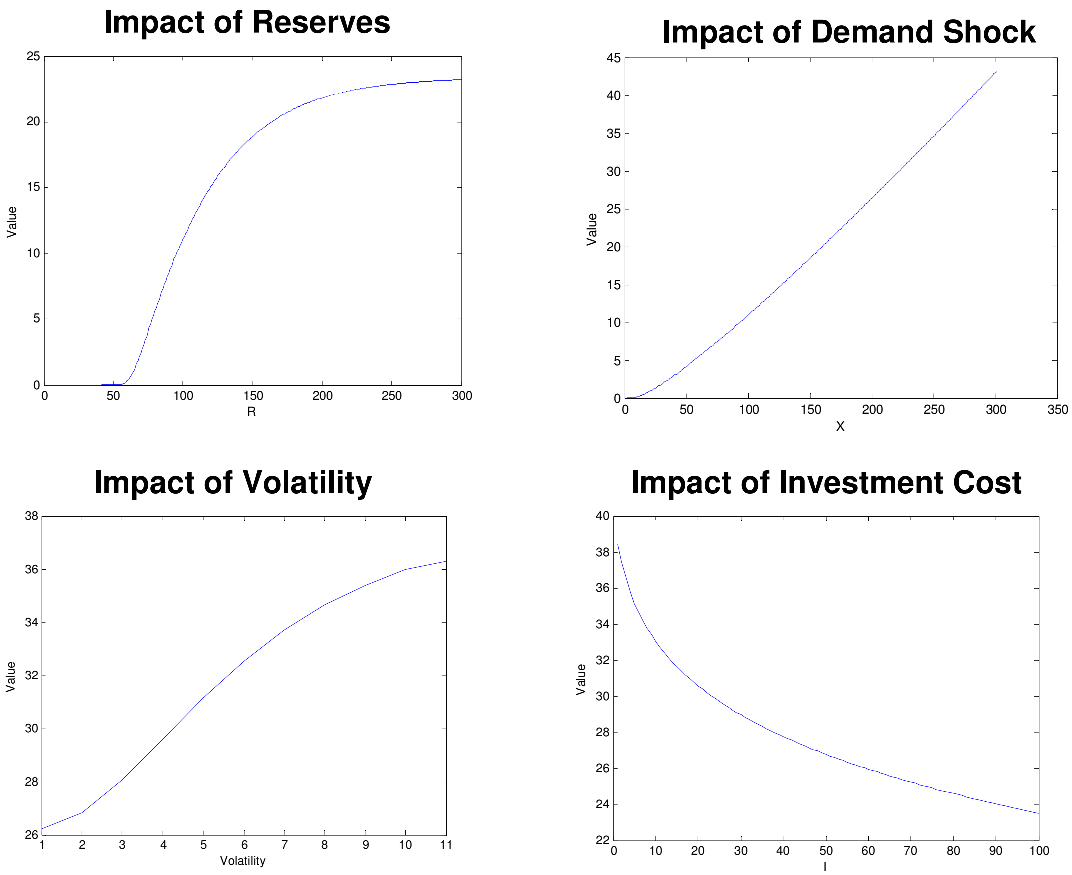



**Figure 6: Value of Capacity Expansion Option**

Up-left panel shows the impact of increasing in the initial reserves on the option value. When R increases the effect of the second state variable is weaken since it now looks more similar to an infinite reserve case and converges to a normal capacity expansion. Down-right panel gives the value of option against its exercise price. The option value tends to zero as the costs to build extra capacity increases which means if it is too costly to add to the capacity the owner treats it as there is no expansion option.

### *3.7) The Exercise Region*

Since the analytical solution for the option value function does not exist the binomial tree method is used to find the optimal trigger curve in the case of isoelastic demand. The interesting feature of this tree is that unlike other capacity choice problems the second state variable (i.e. R) diminishes over time and therefore the current value of option to build capacity deteriorates over time. Since we know that option will only be exercised at the region where the capacity constraint is binding R has a simple deterministic dynamics. Therefore we can use a standard tree for demand uncertainty while changing the value of R according to this dynamics.

Figure 7 shows the trigger curve on the space of R and X.

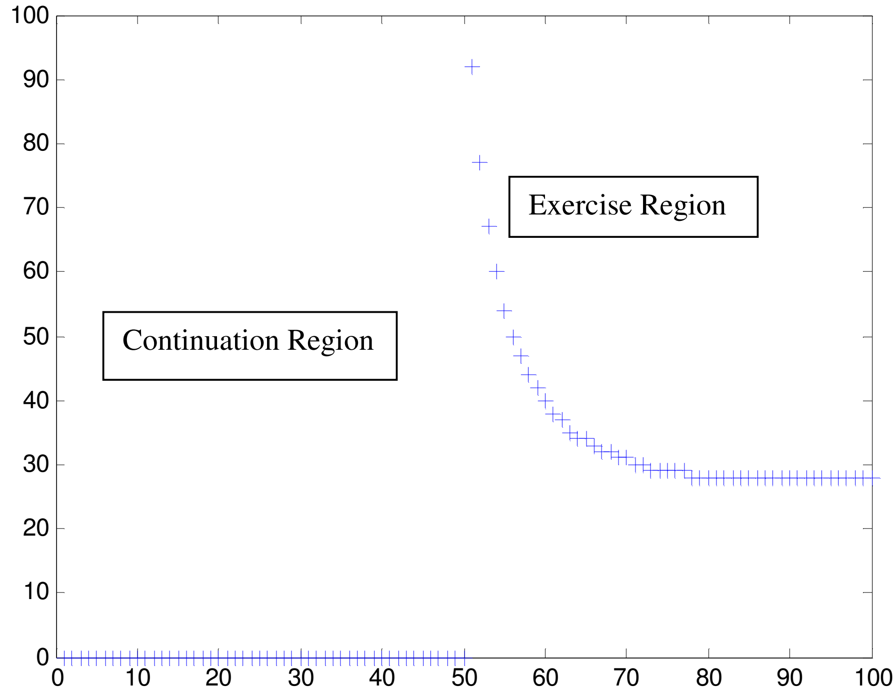



**Figure 7: Exercise and No Exercise Regions for Capacity Option**

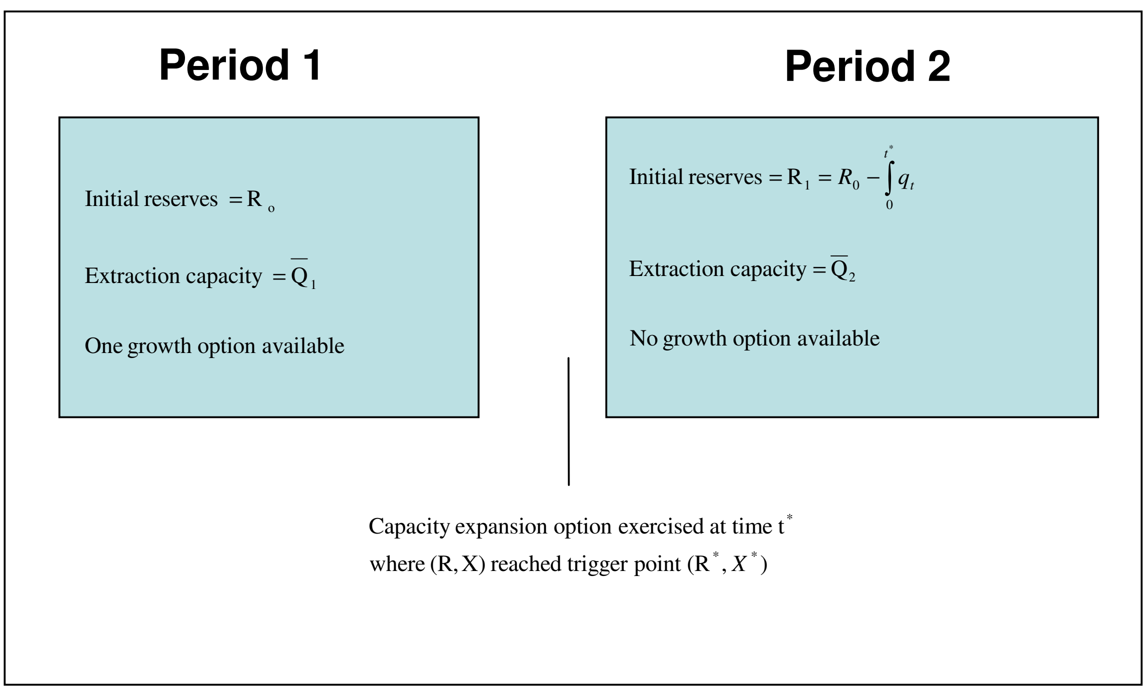
The trigger boundary suggests an intuitive result. With values of R above a certain threshold it is optimal to exercise the option if the demand shock is strong enough. On the other hand if R is below that critical level, the option will never be exercised even with very high demand shock (the cross sign on the horizontal axis specifies values of R where option will never be exercised).

## 4) Optimal Extraction Policy with Capacity Expansion Option and Linear Demand

Unlike isoelastic demand where the realization of demand process does not influence optimal the production rate, under linear demand optimal extraction rate is a function of shocks. In this setting, the optimal trajectory of production is a function of demand shocks and the capacity available. Moreover, as the producer is forward looking she will take into account the possibility of exercising the capacity expansion option in the future. This may change the optimal production rate compared to the scenario where there is no expansion option. This section of the paper aims to investigate this effect first by some simple example and then a numerical simulation.

The overview of producer's two-period problem is demonstrated in the figure 8. Using a backward induction, one sees that the second period problem is a normal extraction problem with given reserves R1 (the amount left from the first period) and the capacity limit $\overline{Q}_2$. The first period problem, however, is more complicated because the producer has a capacity expansion option available and therefore should decide on how much to produce and when to exercise the option. It is a stochastic control problem augmented with an optimal stopping time and non-zero variable terminal value. The producer of the first period takes the fact into account that the value function of second period problem is directly depending on the resources left for the second period. Therefore any policy regarding extraction rate in the first period changes the value of the second period problem and on the other hand any decision to build capacity influences the optimal extraction rate. Therefore, the problem has a fixed-point type behavior.

**Figure 8: Producers Problem**



Denote the value function of the first period by V and that of the second period by W and the time to expand capacity by T. The mathematical formulation of the problem as sum of two depended optimal control problems will be:

$$\underset{\mathrm{T},\, \mathrm{q_t}}{Max} \int_0^T \pi(q_t)e^{-rt}dt + e^{-rT} \int_{\mathrm{T}}^{\infty} \pi(q_t)e^{-r(t-T)}dt$$

$$\dot{R} = -q$$

$$dX = X.\mu.dt + X.\sigma.dW$$

$$R_t, q_t \geq 0$$

$$q_t \leq \overline{Q}_2, \text{if t > T otherwise } q_t \leq \overline{Q}_1$$

The problem would be similar to the constraint case with a major difference that the stopping time will be chosen optimally. This effect change the value function by boundary conditions coming from value matching and smooth pasting conditions.

$$rW = \underset{q \leq Q_2}{Max}\{\pi(q) - qW_R + X\mu W_x + \frac{1}{2}\sigma^2 X^2 W_{XX}\}$$

$$W(0, X) = 0$$

$$W(R,0) = 0$$

$$rV = \underset{q \leq Q_1}{Max}\{\pi(q) - qV_R + X\mu V_x + \frac{1}{2}\sigma^2 X^2 V_{XX}\}$$

$$V(0, X) = 0$$

$$V(R,0) = 0$$

$$V(R, X^*) = W(R, X^*) - I$$

$$V(R^*, X) = W(R^*, X) - I$$

$$V_X(R, X^*) = W_X(R, X^*)$$

$$V_R(R^*, X) = W_R(R^*, X)$$

One should notice that this problem differs from the usual real options and optimal extraction problems because the decisions regarding how much to produce and when to stop and move to the next problem are made simultaneously. In mathematical terms we have a system of partial differential equations with a common free boundary which is more challenging to solve even numerically compared to ODE or single PDE problems.

It is not trivial a priori whether the optimal plan for this case will necessarily differ or will conform to Hotelling's standard solution. To test that first a simple two-period exampled is discussed and then some simulation results will be presented.

### 4.1) A Two-Period Model

Consider a two period problem with stochastic demand shocks and linear demand function. The initial state of shock is $X_1 = 100$ and the demand process in the next period may take two values $X_{2L} = 80, X_{2H} = 200$ with equal probabilities. Maximum production capacity is 50.

The producer solves the following problem:

$$
\begin{aligned}
& Max \pi(q_1) + e^{-0.1} E\pi(q_2) \\
& st \\
& q_1 + q_2 = 100 \\
& q_1, q_1 \geq 0 \\
& q_1, q_1 \leq 50 \\
& \pi(q) = X - q
\end{aligned}
$$

The solution of the unconstrained problem is $q_1 = 34.5, q_{2H} = 64.5, q_{2L} = 40,$ that does not satisfy capacity constraint for high demand realization in the second period. Therefore we move to the corner solution which is $q_1 = 50, q_{2H} = 50, q_{2L} = 40$ and the profit under this production plan is

$$\pi(50,50,40) = (100-50)*50 + e^{-0.1}[0.5(200-50)*50 + 0.5*(80-40)*40] = 6595$$

Notice that with linear demand it is not necessarily optimal to extract all remaining reserves in the last period. Therefore if the demand shock turns out to be low only a proportion of the resources is extracted.

The reason for deviating from the optimal rate of unconstrained problem in the first period is that since the production capacity is limited the producer will not be able to take full advantage of possible high shock. As a result she moves part of the next period production to the first period to use the slack capacity at this period.

Now assume that the producer has a 15 unit capacity expansion option which costs 100 monetary units to exercise. Possessing this option, it is optimal for the producer to wait for a period and then exercise the option only if high shock happens. With this extra possibility the optimal production plan will be $q_1 = 35, q_{2H} = 64, q_{2L} = 40$ that yields a profit of

$$\pi(35,64,40) = (100-35)*35 + e^{-0.1}[0.5(200-64)*64 + 0.5*(80-40)*40] - 100 = 6837 > 6595$$

### 4.2) A Multi-Period Simulation

The results of two-period model are valid for a more general case. Figure 9 depicts the production policy of the producer when the costs of exercising the capacity option is too high and when it is low.

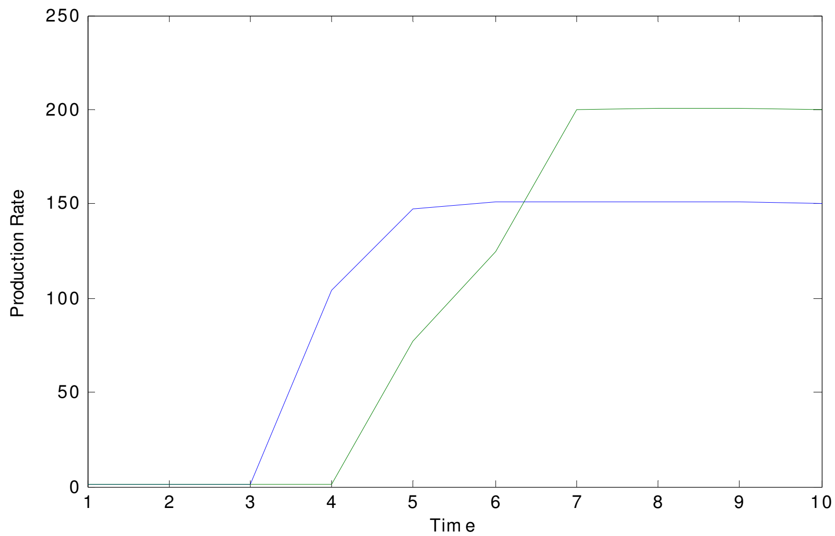



**Figure 9: Optimal Production Rates for Expensive and Low-Cost Options**

The two-period simple model and the multi-period simulations provide some interesting results:

a. The producer with a feasible capacity option(s) chooses a more volatile production plan compared to the producer without the expansion option(s).

b. If there is a high demand shock the producer who does not have any expansion option will response to that shock more strongly (provided that she has enough slack capacity) than a producer with capacity option. The second one may save more resources for some future periods where even higher shocks may take place and use the option there.

These results may justify why the predictions of Hotelling rule are not happening in reality. The rule suggests that the production rate should gradually decrease the price should increase with the interest rate. However, many studies show that the real price of commodity tend to be almost constant during the last century and hardly increased. If one adds the capacity constraint aspect to the problem then having a close to flat production rate will be expected when investment in higher capacity is not feasible.

This analysis is able to show the path-dependent behavior of the capacity expansion option as well. Comparing two different paths for the demand process in past, one can conclude that if there was a strong shock in future the rate of extraction was higher than if there was a low shock. Therefore, the level of reserves today will be different for two producers with two different histories. Now, if both producers face the same demand shock their response in terms of building further capacity might be different since the producer with higher current reserves (lower past production) is more willing to build capacity than a producer who have produced high in the previous periods and therefore retains lower level of reserves.

## *5) Price Dynamics*

For a monopolist the market price highly depends on the production rate. Since the existence of option changes the production rates we expect that the price dynamics will also be different for the producer with and without capacity option. Moreover, one can investigate the impact of considering option value in making investment on price dynamics by comparing the price process in this case with a myopic NPV rule.

For a producer facing isoelastic demand the price dynamics will be given by:

$$P = X.q^{\gamma} \Rightarrow dp_t = dXq^{\gamma} + Xd(q^{\gamma}) = (X\mu dt + X\sigma dW)q^{\gamma} + Xq^{\gamma-1}dq$$

One can distinguish three regions for price dynamics:

1) $\mathrm{q}^* > \bar{q}_1 \Rightarrow q_t = \bar{q}_1 \Rightarrow dq = 0, \mathrm{dp} = (X\mu dt + X\sigma dW)q^{\gamma} = (p\mu dt + p\sigma dW)$

2) $\mathrm{q}^* > \bar{q}_2 \Rightarrow q_t = \bar{q}_2 \Rightarrow dq = 0, \mathrm{dp} = (X\mu dt + X\sigma dW)q^{\gamma} = (p\mu dt + p\sigma dW)$

3) $\mathrm{q}^* < \bar{q}_1 \Rightarrow dq = 0, \mathrm{dp} = (X\mu dt + X\sigma dW)q^{\gamma} + Xq^{\gamma-1}dq = (p\mu dt + p\sigma dW) + Xq^{\gamma-1}q\frac{r-\mu}{\gamma}dt =$

$$p_t[(\mu + \frac{r-\mu}{\gamma})dt + \sigma dW]$$

Region 3 suggests that when the capacity is not binding the drift term of price is larger. This is due to the fact that in the constraint region Hotelling rule can not hold completely and the time dynamics of q which causes change in price is absent.

Now let's look at the price paths a fixed realization of demand shock under real options rule and NPV rule. First we show the case for a producer without capacity option. As expected the price follows the dynamics of X in the binding region and then the slope jumps as Hotelling effect is now in place.

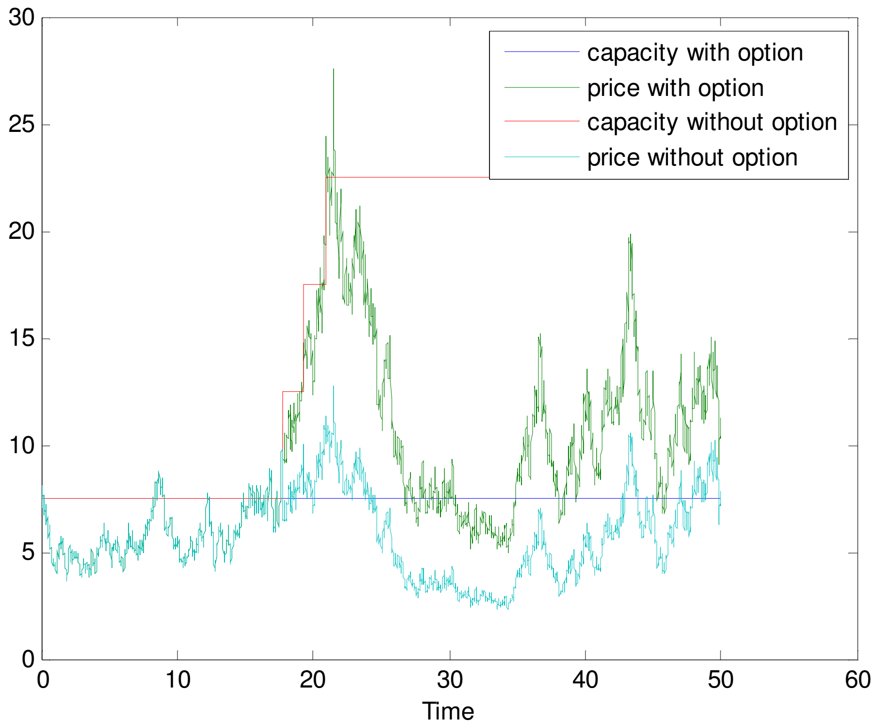



**Figure 9: Behavior of Capacity and Price without and with Option**

In the second scenario we let the producer exercise the capacity option when it is optimal to do so. We see that for low realizations of X the capacity option is not exercised and the price looks like the previous case. But for high values of X the producer adds to the existing capacity and this causes price to drop significantly and then follow the usual dynamics.

One implication of this behavior is that if the reserves are approaching the end of their life the price will be more volatile. Price volatility is caused by the fact that with lower level of reserves remained the probability of capacity expansion option exercised is lower and therefore we expect the supply side to react less aggressively to the demand shocks which cases higher price hikes and more volatility.

## 6) Conclusion and Extensions

In this paper the interaction of capacity building and extraction decisions of an exhaustible resource monopolist was studied. Several results were derived using numerical solutions to the continuous-time optimal control problems. First, it was shown that the capacity expansion capacity for such a firm may have a finite life. The trigger boundary and the sensitivities of option value verses different parameters were illustrated. Finally, it was shown that the existence of the option will change the optimal production policy of the producer. These results have interesting implications for better understanding of the behavior of firms producing exhaustible resources. Unlike the predictions of Hotelling rule we do not necessarily expect the production rate to decline over and the price increase with interest rate because the producer is bounded by the capacity and tries to produce as much as possible if she expects the saved reserves today to be valueless in future. Moreover, prices will be more volatile if the option to build is more expensive and if the remaining reverses are low.

The current research can be extended in several directions. First of all, one can model the problem for a oligopolistic case where two players extract from a common reserve with different capacity limits. Second, new sources of stochasticity can be introduced into the model including jumps in the level of remaining reserves, changes in the production costs (which were assumed to be zero here) and the cost to build more capacity. One interesting variation will also be to look at the results of this paper if the demand process follows a mean reverting process instead of a growing geometric Brownian motion.

## References


- Brennan, M.J. and Schwartz, E.S., 1985, Evaluating Natural Resource Investments, The Journal of Business 58, 2,135-157

- Campbell ,H, 1980, The Effect of Capital Intensity on the Optimal Rate of Extraction of a Mineral Deposit, The Canadian Journal of Economics, Vol. 13, No. 2, pp. 349-356

- Carlson M, Z Khokher, S, Titman, 2007, Equilibrium Exhaustible Resource Price Dynamics, Journal of Finance, vol. 62(4), pages 1663-1703, 08

- Dangl, Thomas, 1999, Investment and capacity choice under uncertain demand, European Journal of Operational Research, 117/3,pp. 1-14

- Davis, R. and D Moore, 1998, Valuing mineral reserves when capacity constrains production, Economics Letters 60 (1998) 121–125

- Dixit, A. and R. Pindyck, 1994, Investment Under Uncertainty. Princeton, NJ: Princeton University Press.

- Hotelling, Harold, 1931, The economics of exhaustible resources, Journal of Political Economy 39, 137–175.

- Litzenberger, Robert H., and Nir Rabinowitz, 1995, Backwardation in oil futures markets: Theory and empirical evidence, Journal of Finance 50, 1517–1545.

- R Morck, E Schwartz and D Stangeland, 1989, The Valuation of Forestry Resources under Stochastic Prices and Inventories, Journal of Financial and Quantitative Analysis, Vol 24, No 4

- McDonald, R., and D. Siegel, 1986, The Value of Waiting to Invest, Quarterly Journal of Economics 101, 4: 707-727.

- Pindyck, Robert S, 1980. Uncertainty and Exhaustible Resource Markets, Journal of Political Economy, University of Chicago Press, vol. 88(6), pages 1203-25, December