INVESTMENT, TAXES, AND UNCERTAINTY,
WITH APPLICATIONS TO THE NORWEGIAN PETROLEUM SECTOR

by

Diderik Lund
Cand. mag., University of Oslo, Norway (1975)
Cand. polit., University of Bergen, Norway (1981)

SUBMITTED TO THE DEPARTMENT OF ECONOMICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

PH.D. IN ECONOMICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
January, 1987

Copyright Diderik Lund, 1987

The author hereby grants to M.I.T. permission to reproduce and
to distribute publicly copies of this thesis document in whole
or in part.

Signature of Author ........................................
Department of Economics
January, 1987

Certified by ...... ........................................
Professor James M. Poterba
Thesis Supervisor

Accepted by ...... ........................................
Professor Richard S. Eckaus
Chairman, Departmental Graduate Committee
Department of Economics
INVESTMENT, TAXES, AND UNCERTAINTY, WITH APPLICATIONS TO THE NORWEGIAN PETROLEUM SECTOR
by
DIDERIK LUND

Submitted to the Department of Economics in January, 1987 in partial fulfillment of the requirements for the Degree of Ph.D. in Economics

ABSTRACT

The thesis consists of four essays, all related to the public management under uncertainty of the petroleum sector in Norway.

The first essay analyzes the country's petroleum tax system. Adopting contingent claims valuation techniques, after-tax values to a U.S. oil company of alternative development plans for a discovered offshore oil field are estimated. Strong distortionary effects compared to a no-tax situation are found.

The second essay derives the investment criterion for an agent who is constrained to hold more of a risky asset than is optimal at the going market price, but who is presented with some limited opportunity of investing in more of the asset at a lower price. A covariance measure of risk is shown to be valid under very weak assumptions.

The third essay discusses a published Norwegian model considering petroleum in a national portfolio context. It is shown that the proposed solution is wrong, and that the problem, the way it has been formulated, hardly has an analytic solution.

The fourth essay considers the general Norwegian corporate income taxation, which makes up part of the petroleum taxes. An anomaly of the system has previously been neglected in the literature. For shareholders in high tax brackets dividends are more leniently taxed than other cash income.

Thesis Supervisor: James M. Poterba
Title: Assistant Professor of Economics
Acknowledgments.

James Poterba has been my first adviser as I worked on this thesis. I owe him deep thanks for his concerned, at times enthusiastic, advice, which reflected both his creativeness and the necessary realism. The circumstances were somewhat special in that I did most of the work as a non-resident, at the University of Oslo, Norway. Jim was helpful in overcoming the difficulties created by this, and gave intensive tuition during the three periods in which I visited M.I.T. while working on the thesis.

When I asked Robert Solow to be my second adviser, I was just about to leave for Norway for the second time, but he did not hesitate long before agreeing. Although engaged at a late stage, Bob has provided a very useful experience, overview, and point of reference.

These two were also my excellent teachers in public economics and natural resource economics, among other things. I would like also to extend my thanks to the other teachers I had at M.I.T., and in particular mention two from whom I learnt material relevant for this thesis, Mervyn King and Robert Merton. Conversations with the latter two were very helpful.

The M.I.T. Energy Laboratory, Center for Energy Policy Research, provided a desk and a stimulating environment. In particular, I would like to thank Morris Adelman, Charles Blitzer, Panos Cavoulakos, David Laughton, Jeff MacKie-Mason and Sayeed Sadeq for many discussions. In Oslo, Steinar Strem, Vidar Christiansen, and Michael Hoel have been good sources of help. I have talked to numerous other people, and extend my thanks to fellow students, teachers, and colleagues at M.I.T. and in Oslo.

A special thank to Petter Laake and Tore Schweder for letting me use their computers every night and weekend for several weeks, and to Inger Johanne Kroksjø for preparing the figures.

I gratefully acknowledge financial assistance from the following Norwegian sources: The Norwegian Research Council for Science and the Humanities, the University of Oslo, the Ministry of Petroleum and Energy, and four foundations, two at the University of Oslo ("Christiania Bank og Kreditkasses legat" and "August og Mathea Mantheys legat"), moreover, the Bank of Norway's foundation for economic research, and "Grinifangers Kulturfond".
TABLE OF CONTENTS:

CHAPTER 1: INTRODUCTION .......................................... 5
References for chapter 1 ........................................... 22

CHAPTER 2: SOME INCENTIVE EFFECTS OF THE NORWEGIAN
PETROLEUM TAX SYSTEM UNDER PRICE UNCERTAINTY,
ANALYZED BY CONTINGENT CLAIMS VALUATION
TECHNIQUES ............................................................ 24
Notes to chapter 2 ..................................................... 81
References for chapter 2 ............................................ 98
Appendix 1 to chapter 2 ............................................. 101
Appendix 2 to chapter 2 ............................................. 113

CHAPTER 3: INVESTING IN NON-MARKETABLE ASSETS .............. 122
Notes to chapter 3 ................................................... 152
References for chapter 3 ........................................... 157

CHAPTER 4: COMMENTS ON THE SOLUTION OF A DYNAMIC
PORTFOLIO MODEL WITH OIL ................................. 159
Notes to chapter 4 ................................................... 189
References for chapter 4 ........................................... 192
Appendix to chapter 4 ............................................. 193

CHAPTER 5: LESS THAN SINGLE DIVIDEND TAXATION ............. 196
Notes to chapter 5 ................................................... 209
References for chapter 5 ........................................... 211
CHAPTER 1.

INTRODUCTION.
1. Background: Norwegian Petroleum.

The first three essays in this thesis were inspired by a set of important problems in the economy of my home country, Norway. Uncertainty is a major concern in the management of the country's petroleum sector, and the methods to cope with it are in rapid development.

The first discovery of petroleum, i.e. crude oil and natural gas, in Norway was made offshore, in the North Sea, in 1969. The first commercial production took place in 1971. By 1985, the sector accounted for 17.3% of GDP, 19.6% of real investment, 36.2% of exports, and 34.8% of taxes. The present fall in the crude oil prices, by about 50% from 1985 to 1986, has reduced these shares, but it has made the problem of uncertainty even more apparent to the public.

The Norwegian experience may be somewhat special, in that Norway is one of the few industrialized nations that has exported most of its extracted petroleum. The Netherlands is another case, but the Dutch have mainly extracted natural gas from one large field, and this poses somewhat different problems. However, some of the Norwegian fields are also very large, and this motivates detailed research into the microeconomic question of how to develop a single field. Whether to install one more deep-water platform at a cost of two billion dollars is not a trivial question.
Uncertainty is encountered by the different agents at various levels of the petroleum activity. The first essay (chapter 2) concerns taxation under uncertainty. In order to analyze this, it is necessary to model the behavior of oil companies under uncertainty. The second essay deals with an investment criterion in a case of incomplete diversification. This criterion is intended for public investment decisions in the Norwegian petroleum sector, at least for decisions concerning the development of single fields. The third essay discusses a model of the overall extraction strategy, in the context of national portfolio planning.

These three topics are interrelated. The investment criterion for the Norwegian government will depend on the share of the value of oil in the national portfolio, which should depend on the extraction strategy. To design a good tax system, one should take into account the way in which the government values its tax claim, which follows from its investment criterion or "discount rate". A bit more about these interrelations will be said in this introductory chapter.

The fourth essay follows from my work on petroleum taxation. A previously neglected anomaly of Norwegian corporate income taxation (which is part of the petroleum taxes) is analyzed.
2. The Tax Model.

There are two branches of the natural resource taxation literature. One considers the intertemporal aspects of taxation, i.e. the introduction of various tax systems into intertemporal models of optimal extraction or of market equilibrium. This literature is summarized in chapter 12 of Dasgupta and Heal (1979).

The other branch is concerned with the more static questions of effects of the tax system on exploration effort and development methods. I believe this branch is known to a more specialized audience, but it has engaged a large number of authors. Central references are the seminal article by Garnaut and Clunies Ross (1975), moreover Garnaut and Clunies Ross (1979) and (1983), Sumner (1978), Dowell (1979), Palmer (1980), Emerson (1980).

Chapter 2 should be seen as an improvement of the methods in this second branch, and the methods applied by governments in evaluating tax reforms. The cited authors are not always explicit when it comes to the modelling of uncertainty and behavior under uncertainty, although they all realize that this is important. Garnaut and Clunies Ross do not include formal models of uncertainty. What they state verbally may be summarized as an assumption that firms are variance-averse. In addition to the cited authors, there are studies that
neglect uncertainty altogether, such as the official studies by the Norwegian Ministry of Finance (1980), (1986).

In chapter 2, the assumption is that the oil company is a U.S. corporation, owned by well-diversified U.S. shareholders. This makes the lessons about valuation from finance theory more easily applicable than if the owner had been some government. Of these lessons, the contingent claims analysis has recently been adopted to value tax claims, and this method is applied in the present study.

Contingent claims analysis, also called option valuation, gives the value under uncertainty of a claim which is a non-linear function of some underlying uncertain variable. The oil prices are the uncertain variables in chapter 2. Of course there are other uncertain variables facing an oil company, but at this stage of our knowledge, a simplification was felt necessary.

The method is particularly relevant when the tax claim is asymmetric, in the sense that profits are taxed at a high rate, whereas losses are not refunded, at least not at the same high rate. When there is a significant probability of losses, the tax claim has a value under uncertainty which is higher than its value under full certainty if the oil price is set to the expected oil price. This point has been recognized
before, e.g. by Mayo (1979). The improvement offered by contingent claims analysis is that a number for the value can be found.

The petroleum taxation in Norway is a very complicated function of oil prices. Closed form solutions to the contingent claims values cannot be found. Chapter 2 values the tax claim under specific assumptions, using Monte Carlo simulation. This simulation technique is an implementation of a result by Cox and Ross (1976) about contingent claims values.

The Monte Carlo simulation gives an estimate of the value of the tax claim under assumptions of continuous trading in financial markets, no riskless arbitrage opportunities, the existence of appropriate financial assets (diversification opportunities), and a particular stochastic process for the oil price. As described by Cox, Ingersoll and Ross (1985, p. 380), "the equilibrium price of a claim is given by its expected discounted value, with discounting done at the risk free rate, when the expectation is taken with respect to a risk-adjusted process for wealth and the state variables. The risk adjustment is accomplished by reducing the drift of each underlying variable by the corresponding factor risk premium."

Before the contingent claims analysis had become known, Bøhren
and Schilbred (1980) used Monte Carlo simulation to study the properties of the Norwegian petroleum taxation under uncertainty. They compare stochastic, multi-period cash flows on the basis of their respective net present values, using a constant discount rate. In spite of similarities, the contingent claims analysis in chapter 2 has a very different theoretical basis from their study. They do not suggest why any investor would actually value the claim to the cash flow in the way they do, and they do not adjust the drift of the oil price according to its risk. They acknowledge that their approach can be criticized on theoretical grounds.

The second major improvement in the present study compared to many previous studies, is the modelling of the production possibilities. Some studies, notably Norwegian Ministry of Finance (1980), (1986), do not consider the effect of taxes on production choices explicitly. Without some description of production possibilities, it is impossible to quantify the distortive effects of taxation. The present study focuses on the effects on the scale of development of a discovered oil field. This has been studied by other authors, by describing a number of development plans for the field, and the resulting time paths of extracted quantities.

Data for the development alternatives that are considered by companies are not in the public domain. Nystad (1985) uses a
relatively complicated model to describe different possibilities, but the model and the numbers that go into it are nevertheless a result of educated guessing only. Such a complicated production model, with dated inputs and outputs, is not easily comprehensible, to economists, to engineers, or to anyone else.

The present study uses instead an analytical production function which is easily comprehensible, although it may be somewhat less realistic than Nystad's. By varying the elasticity of the production function, it turns out that the distortive effects of the tax system are strong irrespective of the elasticity, at least within the range that is covered.

Compared to the existing literature, these are the two major improvements that would be applicable to any tax system. In addition, the present study contains a relatively detailed discussion of the taxation and regulation of companies in the Norwegian petroleum sector. The intention has been that of these rules, no aspects of major economic importance should be neglected, and simplifying assumptions should be explicit, although perhaps expressed in a condensed way.

The main result of the analysis is that the Norwegian petroleum taxation may have strong distortionary effects. Under the assumptions that are made, the companies will want
to reduce the costs of development of a field by more than 50% compared to the before-tax optimum. The total value of the field is thereby reduced by 30 to 40%.

In the analysis in chapter 2, the petroleum tax system is kept fixed under all petroleum price paths. As this study was being finished, the Norwegian Ministry of Finance (1986) proposed a new petroleum tax system as a response to the recent decrease in crude oil prices. It seems, therefore, that the assumption is not realistic.

However, the assumption may be necessary as a part of the process of analyzing the petroleum tax system. The changes in the tax systems will have incentive effects themselves, by creating what the companies regard as political risk, and it would be a good thing if one could construct tax systems that did not have to be changed as prices change. In order to analyze such reforms, one needs methods for analyzing tax systems as if they do not change. Moreover, the assumption of a fixed tax system is prevalent in the literature. In spite of the possible shortcoming at this point, the methods presented here do represent improvements.

Another limitation of the analysis should be made clear. The tax claim is valued as seen from the point of view of a U.S. oil company. The Norwegian government may value the tax claim
differently, as implied by the analysis in chapter 3. This does not alter the validity of the results concerning the effects of the tax system on the behavior of the oil companies. On the other hand, the value to the Norwegian government of its tax claim may not be maximized by the development plan which maximizes the tax claim as valued by the companies. In particular, the revenue-neutral reforms of chapter 2 have no meaning when the government's valuation is different.

Despite the limitations of the analysis, chapter 2 should represent an improvement compared to most existing analyses of petroleum tax systems under uncertainty.

3. The Investment Criterion.

The study by Arrow and Lind (1970) concluded that the public sector should behave as risk neutral, but this conclusion was based on the assumption that the benefits and costs from the projects are uncorrelated with aggregate consumption or other macroeconomic target variables. This does not hold for projects in the relatively large Norwegian petroleum sector.

Models by Diamond (1967) and Sandmo (1972) establish conditions under which a social planner can allocate risks no better than a competitive stock market. Under those conditions, the implication is that the planner should observe
required rates of returns (of companies, or preferrably of single projects) in the stock market, and apply them to similar projects in the public sector. The conditions are very strong. The main problem in applying their models to the Norwegian petroleum sector is that neither Norway as a nation nor most Norwegian citizens adjust their holdings of petroleum optimally, as is the assumptions in theories of stock market allocations of risk. Most Norwegians do not own shares in oil companies, but choose a corner solution. Short-selling is not legal. They hold petroleum implicitly, as claims to future government services to be financed by the government's claim on petroleum. For corner solutions, the standard results are not valid.

One could imagine two situations in which a government could observe required expected returns on oil projects. One would be an internal stock market, where the petroleum wealth of a country was traded, and the citizens' required expected returns could be observed. These would certainly reflect the preferences, at least of those citizens that participated in the trading, and did not choose a corner solution. The corner solution is the problem in the Norwegian context.

The other would be an international stock market, in which the government or the citizens participated. In order for the required expected returns in the international market to
reflect the preferences of the citizens, the nation would have
to be well diversified internationally. If all oil rich
nations were well diversified internationally, there would in
effect be no distinction between oil nations and non-oil
nations from a portfolio point of view, only physically. Oil
nations would sell claims to oil in their ground.

This contradicts the facts. It seems more realistic to
describe oil in the ground as an asset that can not be traded
internationally, or only to a modest extent. The fear of
nationalization, regulation, and/or taxation prevents
diversification.

Lind (1982) and Adelman (1986) discuss the choice of a social
discount rate for risky investments in petroleum projects, or
more generally, energy projects. Both of these apply the
lessons from the Capital Asset Pricing Model (CAPM) that
projects should be valued according to their systematic risks,
i.e. the project returns' covariances with the returns on a
market portfolio, or on national wealth. It is a problem that
the CAPM, and most extensions of it, assume that investors
hold well diversified portfolios. Adelman discusses oil
exporting LDCs in particular, and he is aware that these
countries are not well diversified.

A central question is therefore what an investment criterion
under uncertainty looks like when the investor is imperfectly diversified (in some sense). The model in chapter 3 shows that the criterion is very similar to the ones that are known from the CAPM literature. This gives a theoretical fundament for applying CAPM type reasoning to non-diversified investors. Moreover, the essay brings out the generality of the covariance-based risk concept, by discussing the exact conditions under which it may be used.

The general measure of investment risk is the covariance of the project return with marginal utility of consumption. This does not depend on an assumption that utility can be expressed as a function of only the mean and the variance of consumption or wealth. With this latter assumption, however, the correct risk measure becomes the covariance of the project return with the investor's wealth. For an imperfectly diversified investor, the risk measure can not be expressed in terms of the covariance with variables that are aggregated over all investors, except for project returns that are linear combinations of returns of marketable assets.

Risk should be measured by a covariance measure even if the investor is unable in any way to compose an optimal portfolio of risky assets. In order to base the risk measure on the covariance with the marginal utility of consumption, the requirement in a multi-period model is that consumption is
chosen optimally.

The application to oil is not quite obvious, since a nation which has oil, does not invest in oil in the sense of buying more of it. However, from the point in time of investment in an offshore oil platform, until oil is extracted, the systematic risk involved is exactly the oil price risk. And at the margin, when considering how much to extract from a given field, the question is actually how much oil to hold in the portfolio during the period between investment and extraction, since the oil which is left in the field is most often lost for economic purposes.

In order to derive a criterion for public investments from the model in chapter 3, one must be willing to postulate that society has a risk-averse expected utility function, which can best be thought of as the utility function of a representative individual. This is of course less satisfactory than the approach of Diamond and Sandmo. Their models do not postulate particular social welfare or utility functions, but develop the investment criterion from much weaker conditions in this respect.

However, postulating a risk-averse expected utility function for society seems better, in my judgment, than having a model which neglects particular facts, such as the correlation
between the oil price and Norwegian national wealth, or the lack of diversification in the Norwegian national portfolio.

I believe the results from chapter 3 has a much more general application, too. The economic literature knows many studies that have assumed that a firm, or a decision maker, has the expected utility of profits as the objective function. One example is Sandmo (1971). Related to natural resource extraction is Sundaresan (1984). Related to taxation of petroleum are Black et. al. (1982), and Garnaut and Clunies Ross (1979), who write, "we suppose risk aversion on the part of the investor. When an investor is risk-averse, a project the outcomes of which have greater variance is less attractive to him than one whose outcomes have lesser variance and the same expected net present value (when discounted at the same rate)," (p. 195).

This assumption is clearly at variance with the lessons from the CAPM literature, given that the investing firm is owned by well diversified shareholders. But even when it is not, I show in chapter 3 that the investor should consider a covariance measure, not a variance measure, in order to determine the relevant risk of a project. Only when the investor has no other sources to derive utility from, the two become the same.
4. Oil in a National Portfolio Model.

At the macroeconomic level, the Norwegian petroleum sector has raised the problem of national portfolio planning, and the optimal extraction strategy under uncertainty. The most widely read study of this, Aslaksen and Bjerkholt (1986), is recently published in a book at the M.I.T. Press. It has also appeared in two other books, Aslaksen and Bjerkholt (1985a) and (1985b).

Chapter 4 is a thorough discussion of their model, both the definition of the problem, various assumptions, and the proposed solution. I show that based on the way they set up the problem, it has no analytical solution. Their proposed solution is based on at least three different mathematical mistakes.

Of their assumptions, I pay particular attention to the stochastic properties of the oil price, to the ability of the planner to sell and buy oil, and to the utility function. The oil price is assumed to follow an exponential path with additive noise. This is not a common assumption. In particular, it is quite different from the price process that is found in chapter 2, which in discrete time is a geometric random walk, although the expected price path as of time zero is the same. Combining these two alternative assumptions with two alternative assumptions on the planner's ability to buy
and sell oil, I show that the sign of partial effect of a change in the oil price will depend heavily on the choice of assumptions. More research is needed on the stochastic properties of the oil price.

In commenting on the same model, Newbery (1986) writes that "Few would disagree that the problem is important, but most economists and policy-makers might be tempted to throw up their hands and claim that the problem is too difficult for quantitative analysis," (p. 195). In accordance with this, I have not been able to come up with an alternative model which has an analytic solution.


The framework of King (1977) is very useful for the analysis of corporate income taxes. King finds that among the 19 countries he looks at, Norway has the corporate income tax which is the least discriminatory against dividends compared to other cash income. He finds no such discrimination at all, that is "single" as opposed to other countries' various degrees of "double" taxation of dividends. However, it is shown in chapter 5 that King misses an important element of the Norwegian system, that a generalization of King's categories is needed to capture this element, and that Norway actually has "less than single" taxation of dividends for individuals in high tax brackets.
References.


Garnaut, Ross, and Anthony Clunies Ross (1975), "Uncertainty, Risk Aversion, and the Taxing of Natural Resource
Garnaut, Ross, and Anthony I. Clunies Ross (1979), "The
Neutrality of the Resource Rent Tax," Economic Record,
vol. 55, no. 150, pp. 193 - 201.
Garnaut, Ross, and Anthony Clunies Ross (1983), Taxation of
King, Mervyn A. (1977), Public Policy and the Corporation,
to the Discount Rate for Evaluating National Energy
Options," in Robert C. Lind (ed.), Discounting for Time
and Risk in Energy Policy, Baltimore, Maryland: The Johns
Hopkins University Press, pp. 21 - 94.
55, no. 150, pp. 202 - 213.
Newbery, David M. G. (1986), "Comment," in J. Peter Neary and
Sweder van Wijnbergen (eds.), Natural Resources and the
195 - 200.
Norwegian Ministry of Finance (1980), "Ot. prp. nr. 37, (1979-
80)", (Bill to the Norwegian Parliament,) Oslo, Norway.
Norwegian Ministry of Finance (1986), "Ot. prp. nr. 3, (1986-
87)", (Bill to the Norwegian Parliament,) Oslo, Norway.
Palmer, Keith F. (1980), "Mineral Taxation Policies in
Developing Countries: An Application of Resource Rent
27, no. 3, pp. 517 - 542.
61, no. 1, pp. 65 - 73.
Sandmo, Agnar (1972), "Discount Rates for Public Investment
13, pp. 287 - 302.
Sumner, M. T. (1978), "Progressive Taxation of Natural
Sundaresan, Suresh M. (1984), "Equilibrium Valuation of
Natural Resources," Journal of Business, vol. 57, no. 4,
pp. 493 - 518.
CHAPTER 2.

SOME INCENTIVE EFFECTS OF THE NORWEGIAN PETROLEUM TAX SYSTEM
UNDER PRICE UNCERTAINTY,
ANALYZED BY CONTINGENT CLAIMS VALUATION TECHNIQUES.
1. **Introduction.**

Many countries have introduced special royalty and/or tax systems in order to secure that a large share of the realized rent from the extraction of natural resources is channeled to the treasury. The exploration and extraction are typically carried out by companies in the domestic private sector, or by foreign companies. High tax rates will in these cases often lead to strong distortionary effects.

In order finally to obtain crude oil from the ground, companies make a series of decisions. For simplicity, consider the following five groups: 1. Exploration decisions. 2. Time for development of a discovered field. 3. Scale of development of a discovered field. 4. Operating decisions. 5. Abandonment decision(s).

This study will concentrate on type 3 in isolation. Since the different decisions depend on each other in complex ways, in particular under price uncertainty, I need to justify, at least for the Norwegian deep-water offshore oil fields, why it is not unreasonable to consider the scale decision by itself. First, when development plans are made, exploration is sunk cost, and should therefore be disregarded. Secondly, for any chosen time to start development, the companies will want to develop in a way that is optimal for them at that time. It is thus justifiable to disregard decisions in group 1 and 2.
However, in principle the groups 4 and 5 should be considered when an optimal scale decision is made, since operation and abandonment follow development in time. I assume that the operating flexibility can be neglected because under the existing technology, companies can be expected to operate developed deep-water fields in ways which do not depend much on realized prices. Abandonment flexibility exists, however, and is neglected mainly because it would complicate the model very much, whereas its economic significance is believed to be relatively small.

There are also reasons why the scale decision is one of the most interesting in connection with incentive effects of taxes. At least in the Norwegian context, and I believe elsewhere as well, the time for development start is regulated more directly by the authorities than is the scale. This is done in part by the licensing policy of the government, by which it distributes new, unexplored areas. In Norway the government has additional legal means for regulating the time for development after a field has been discovered.

When it comes to methods for development, however, like methods for exploration, there is an information asymmetry. The motive for letting private companies into government areas, or foreign companies into the nation, is exactly that
they have un-marketed knowledge that they are expected to apply as they strive for profits. Their incentives to do so will be important. The government is unable to detect whether the companies choose a development solution (among alternatives that are unknown to the government) that is in the government's or nation's interest. The choice will be modelled as one-dimensional in this study - a choice of scale.

Tax effects on the choice of scale have previously been studied by Odell and Rosing (1976), Asheim and Furset (1982), and Nystad (1985). The first of these studied the British petroleum tax system, the second studied the Norwegian, while the third studied a hypothetical, stylized tax system. They all used one or more particular examples of field production possibilities, describing what output will be for a limited number of platform configurations. They all agree that petroleum fields are very different, but they claim that the results from their particular studies nevertheless tell something about the effect in general of petroleum taxation on extraction. None of the three studies took uncertainty into consideration in any formal way.

The present study will look at the existing Norwegian petroleum tax system, in more detail than did Asheim and Furset (1982), and will use an analytical production function which describes the production possibilities for the field.
Option valuation techniques, also called contingent claims analysis, are used to analyze the effects of the tax system under uncertainty.

Based on the analytical production function, I can describe alternative development plans at various scales for the same oil field. By calculating the value to the company of undertaking each of these plans, I find the plan which will be preferred. This procedure will be repeated under different assumptions on the company's tax position, the convenience yield on oil, and on the production function elasticity.


Formally, there are only some minor, negligible fees to be paid in order to get a Norwegian petroleum license. (What governs the actual distribution of licenses to companies is another question, which will be ignored here.) The taxes and the state participation are the major tools by which the government takes its part of the rent. 6

2.1. The Formal Tax System. 7

There is a royalty, \( t_r \), levied on the gross value of extracted petroleum. This is a flat 10% on licenses granted before 1972. After 1972, it is 12.5% flat on natural gas, but a progressive scale from 8% to 16% on oil, depending on the monthly production from a field. 8 Apart from the royalty, all
taxes are levied on a company basis. All my calculations will consider one company only. In this study I will consider companies which have offshore petroleum activities only.

Apart from the royalty, there are three main taxes, and two additional ones. The taxes are, with current rates in parentheses:

1) The royalty, at rate $t_r$. $(0.08 - 0.16)$
2) The "income" tax, at rate $t_y$. $(0.23)$
3) The corporate state tax, at rate $t_s$. $(0.278)$
4) The special petroleum tax, at rate $t_x$. $(0.35)$
5) The withholding, or source, tax, at rate $t_w$. $(0.15)$
6) The capital tax, at rate $t_k$. $(0.004)$

The three main taxes, 2), 3), and 4), are not deductible in each other's bases. This gives an overall tax rate of 85.8%, neglecting the royalty and neglecting the timing of taxes. The royalty is deductible in the bases for taxes 2), 3), and 4). Distributed dividends are partly deductible in the base for tax 3).

For simplicity the formulae below will include royalty at a single rate, and a single petroleum price. If the profits before tax ($PBT_{t}$) is positive, i.e. if

$$(1) \quad PBT_{t} = (1-t_r)PRI_{t}QUA_{t}-OPC_{t}-NIE_{t}-DPR_{t}-LOS_{t-1} > 0,$$
then the taxes accrued in year $t$ are:

\[
(t_y + t_s + t_x) \left[ (1-t_r) \text{PRI}_t \times \text{QUA}_t - \text{OPC}_t - \text{NIE}_t - \text{DPR}_t - \text{LOS}_{t-1} \right] \\
+ t_r \text{PRI}_t \times \text{QUA}_t - (t_s - t_w) \text{DIV}_t + t_w \text{NDD}_t - t_x \text{UPL}_t + t_k (\text{BVA}_t + \text{CSH}_t - \text{DEB}_t)
\]

where $\text{PRI}_t$ = petroleum price in year $t$,
$\text{QUA}_t$ = extracted quantity in year $t$,
$\text{OPC}_t$ = operating costs in year $t$,
$\text{NIE}_t$ = net interest expenditures in year $t$,
$\text{DPR}_t$ = depreciation allowance in year $t$,
$\text{LOS}_{t-1}$ = loss carried forward from year $t-1$,
$\text{DIV}_t$ = deductible dividends in year $t$,
$\text{NDD}_t$ = non-deductible dividends in year $t$,
$\text{UPL}_t$ = "uplift" allowance in year $t$,
$\text{BVA}_t$ = book value of fixed assets in year $t$,
$\text{CSH}_t$ = financial assets in year $t$,
$\text{DEB}_t$ = debt in year $t$.

In this case, $\text{LOS}_t$ is zero. (Of the variables above, only $\text{NIE}_t$ is allowed to be negative.)

If there is not a taxable profit, dividends are by law zero, whereas the royalty and the capital tax still accrue. The accumulated tax deductible loss to carry forward will be
(3)  \[ \text{LOS}_t = \text{LOS}_{t-1} - [(1-t_r)\text{PRI}_t \cdot \text{QUA}_t \cdot \text{OPC}_t \cdot \text{NIE}_t \cdot \text{DPR}_t]. \]

This completes the description of the formal tax system. But there are two additional arrangements that should be mentioned. One is the state participation scheme, which may be seen as a means of taxation. The other is the required build-up of a reserve fund within the company, which affects the financial decisions.

2.2. State Participation.

In all licenses granted after 1973, Statoil (or the state directly, which will be seen here as equivalent to Statoil) has had a 50% share or more. In most of these licenses Statoil enjoys a carried interest arrangement in the exploration phase. The other licensees have to pay all exploration costs, and Statoil does not have to refund these, even if a field is discovered. Only then will Statoil enter as a participant and share subsequent costs accordingly.

If it had not been for the carried interest during exploration, the state participation could have been regarded as a kind of cash-flow tax. Such taxes are known to be neutral, see Brown (1948), Ball and Bowers (1982), and Garnaut and Clunies Ross (1983), and will most likely not affect company incentives even under uncertainty. The effect is to
reduce all cash flows (before other taxes) proportionately, immediately.

Of course, the carried interest during exploration is an important disincentive to explore. However, this is in part remedied by a required work program, according to which the licensees must carry out a specific amount of exploration in the license area. The carried interest is still likely to reduce exploration efforts, and, whether there is a work program or not, to reduce the companies' interest in applying for licenses. This is the case even though the exploration costs are tax deductible.

The incentive effect of the carried interest will not be investigated in this study. I will look at the decision on how to develop a field when exploration is sunk cost. However, in discussions of the tax system, one should bear in mind that the companies must be allowed to earn a pure rent from the development and extraction phases as seen separately from exploration, in order to compensate them for exploration. When exploration is sunk cost, this requirement should not affect their decisions on how to develop fields. Ex post, after exploration, any pure rent greater than zero is sufficient for the development to be undertaken. A very low rent will only partly cover the exploration costs, but is better than no rent. But ex ante, in order for exploration to
be undertaken, the companies must expect there to be some compensation to be earned from successful discoveries. If the tax system reduces the after-tax value of the development + extraction phase to approximately zero, no company would be interested in exploration. The requirement ex ante of an expected rent from development + extraction will be significantly higher (at least twice as high) because of the carried interest.

2.3. Reserve Fund Requirement.

The companies are required to build a "reserve fund" of retained earnings.\(^{13}\) The purpose seems to be in part to protect the company's creditors, in part to protect small shareholders. The fund must accumulate until it is equal to at least 20% of the capital stock (\(KST_t\)) (the book value of shares), and equal to at least the difference between the company's debt (\(DEB_t\)) and the capital stock,\(^{14}\) when this difference is positive. This means that until the reserve fund (\(RSF_t\)) fulfills

\[
(4) \quad RSF_t \geq \max\{0.2KST_t, DEB_t - KST_t\},
\]

it must accumulate according to specific rules.

The minimum accumulation at the end of a year depends on the profits and the dividends that year. 10% of profits (after
tax), if positive, must be allocated to the reserve fund. This is called ordinary reserve fund allocation ($\text{ORA}_t$). If dividends exceed 10% of the beginning-of-year total value of capital stock and reserve fund, then an amount equal to the excess dividends must be allocated to the reserve fund. This is called extraordinary reserve fund allocation ($\text{EOA}_t$). This means that if (4) is not fulfilled, then

\begin{equation}
\text{(5)} \quad \text{ORA}_t = 0.1 \text{Max}[	ext{After-tax profits}, 0].
\end{equation}

Moreover, if (4) is not fulfilled,

\begin{equation}
\text{(6)} \quad \text{EOA}_t = \text{Max}[	ext{DIV}_t - 0.1(KST_{t-1} + \text{RSF}_{t-1}), 0].
\end{equation}

If (4) is fulfilled with inequality, then 20% of the excess reserve fund may be distributed as dividends. Dividends in excess of $\text{PBT}_t$ are not deductible for the state tax.

Companies in the Norwegian petroleum sector have to a large extent been debt financed because of the generous interest deductions. (Interest is currently deductible against 85.8% tax.) This has in some cases caused tax authorities to refuse to accept "thin capitalization", i.e. a very high debt/equity ratio.\footnote{15} However, no exact limit has been set. In line with other research on this tax system, I will assume that there exists such a regulation. If the maximum debt/equity ratio is
DEM, then

\[ \text{DEB}_t \leq \text{DEM}(\text{KST}_t + \text{RSF}_t + \text{FRE}_t - \text{LOS}_t), \]

where \( \text{FRE}_t \) is free equity at time \( t \), i.e. non-blocked retained earnings. (For the tax purpose of loss carry-forward, accumulated tax-deductible losses, \( \text{LOS}_t \), are accounted for separately.)

This completes the description of the taxation and related regulations of companies in the Norwegian petroleum sector.

2.4. The Company’s Financial Decisions

There are two reasons why the companies’ financial decisions need to be modeled carefully in this study. First, all previous studies of the Norwegian petroleum taxation have found that the effects of the system are highly sensitive to the choice of debt/equity ratio for the companies.

Secondly, the effect of the reserve fund has sometimes been neglected in previous studies, probably because it is quite complicated, but it is quantitatively significant because there will be large blocked funds within the companies for long periods of their lives. The present study will therefore attempt to model the financial decisions consistently, even though some simplifying assumptions have to be made.
A typical time development of the first petroleum activity of a company will be like this:

1) The company participates in petroleum exploration, and the exploration costs are expensed.20

2) Petroleum is discovered. A field is developed. Losses are carried forward, but development costs are capitalized.

3) The first petroleum is extracted. Losses which have been carried forward, and depreciation allowances, are so high that the first (few) year(s), the company pays only royalties and capital tax, no other taxes.

4) The first taxable profit appears. The company pays all kinds of taxes, distributes dividends, and starts building a reserve fund.

5) Sooner or later the reserve fund catches up with debt minus capital stock, and no more reserve fund allocations need to be made (for some time).

6) As the book value of assets decreases, then so do both debt and the capital stock. The reserve fund becomes too large in relation to these, and it may be decreased each year by 20% of the excess part.

7) If the company starts activities in other fields, it may need to build the reserve fund again.

This is only one likely history. In the present study, investment and operating costs, as well as extracted quantity
and a stochastic process for its price, will be specified for each development plan. The financing of the company will be determined each year by a financial policy. Preferably, an optimal policy should be derived. It is useful first to simplify, neglecting the reserve fund requirement. Assume that the company is in tax position, i.e. no losses are carried forward, such that deduction for interest expenditures and for dividends can be earned immediately.

The company has three possible sources of funds, namely new share issues, retained earnings, and debt. King (1977) develops a method to analyze which is the optimal source of funds. The method is based on assumptions of full certainty, fixed prices, and of taxes being linear. Under those assumptions, a company will in general be able to increase its market value by increasing one source of funds, decreasing the others, and this will go on indefinitely. In reality, however, there will be constraints, e.g. on increasing debt in order to pay out dividends. The method nevertheless tells which is the preferred source of funds up to points where constraints become effective, or the linearity assumptions are violated.

Assume that the shareholders pay a tax rate $t_p$ on both interest and dividend income, and an effective rate $t_g$ on
capital gains. Assume first that \( t_p = 0.3 \), and that \( t_g = 0.1 \). The three relevant parameters of the tax system that should be compared are (current values in parentheses):

\[
\frac{1-t_p}{1+t_w-t_s-t_p} \quad (0.8464),
\]

which gives the tax effect on new share issues,

\[
1-t_g \quad (0.9),
\]

which gives the tax effect on retained earnings, and

\[
\frac{1-t_p}{1-t_y-t_s-t_x} \quad (4.9296),
\]

which gives the tax effect on debt.

The company will prefer the source of financing which has the highest of these parameters. The very generous interest deductions, at a tax rate of 85.8%, clearly implies that debt finance will be preferred. Retained earnings are preferred to new share issues. But the latter conclusion will be reversed for some \((t_p, t_g)\) combinations, approximately when \( t_p < t_g + 0.15 \).

In order to determine the optimal financial policy, it is necessary to determine who are the shareholders of the
company, and who are the lenders. One reason for this is to get some idea about the tax rates of the shareholders, $t_p$ and $t_g$. But in addition, there is the traditional question of non-tax reasons to have a target debt-equity ratio.

The assumption in this study is that the shares are owned 100% by some foreign parent company. This is the reality for a large fraction of the companies. Attention here will be restricted to U.S. owners. The practice is often to borrow money from an affiliate of the same company. This means in effect that all claims to debt and equity are owned by the same parent. The reason why some of it is called debt, is that this is advantageous for tax purposes.

The normal reasons why a company does not choose a corner solution to financing, reasons such as bankruptcy costs and agency problems, do not apply to the subsidiary in Norway in this situation. The debt/equity ratio does not influence the risk to the owners of the two, simply because they are owned by the same entity.

There are two reasons why the parameters of King (1977) are not sufficient to derive the optimal financial decisions of a company in the Norwegian petroleum sector. The first is the legal constraints on dividends, retentions, capital stock, and debt. The other is the uncertainty, in particular about
future profits and about the tax system in the future.

To prove that some financial policy is optimal under such circumstances is a very complicated task. It will not be attempted here. Instead, a set of assumptions on financial policy will be made, and it will be suggested why these look good for the company in most respects.

The company’s balance sheet has the form,

$$\text{BVA}_t + \text{CSH}_t = \text{DEB}_t + \text{KST}_t + \text{RSF}_t + \text{FRE}_t - \text{LOS}_t.$$  \hspace{1cm} (8)

For each development plan to be analyzed, the time sequence of \text{BVA}_t’s is given. At the end of year \(t\), for a given history, \text{LOS}_t is given, either zero or equal to year \(t\)’s accumulated loss. The size of the reserve fund, \text{RSF}_t, is regulated. The company is free to choose \text{CSH}_t, \text{DEB}_t, \text{KST}_t, \text{RSF}_t, \text{and} \text{FRE}_t \text{within the constraint of (8) and some other limitations. One limitation is that none of the variables can take negative values.}

As was noted above, debt finance is very clearly preferred in a case of full certainty. This is so clear, that I will assume that it carries over to the case of uncertainty which I consider here. It follows that the company does not want to hold financial assets, \text{CSH}_t. I assume that the company can
increase or reduce its capital stock, $KST_t$, at will. There is a question whether the company will want to raise the equity required by (7) by issuing shares or by retaining earnings. The parameters that would determine this under full certainty, are very close in value, and depend on the personal tax rates of the shareholders.

However, there is a good reason not to retain earnings apart from the required reserve fund. Dividends paid out of the year's profit are deductible in the base for the state tax, currently at 27.8%, and a retention for later distribution would not earn this deduction. The parameters of King (1977) can be adjusted to take this into account, and the desirability of not retaining earnings becomes very clear. I conclude by assuming that the companies never retain free equity, i.e. $FRE_t$ is always zero.

As long as the company can avoid holding positive net financial assets, it will choose levels of $KST_t$ and $DEB_t$ that make (7) satisfied with equality. The exception is when $RSF_t$ is too large to make this possible. Then $KST_t$ is zero, and $DEB_t$ is determined as a residual.

If the reserve fund exceeds $(BVA_t - LOS_t)$, debt is zero, and the company will be forced to hold financial assets, under the assumption that it does not invest in new real assets.
Since the return on the financial assets is taxed at 85.8% currently, it is not desirable to hold them, and the company will want to avoid coming in a situation with such a large reserve fund. It will reduce the fund whenever a reduction is legal. Apart from the legal 20% reduction of any excess over the target, given in (4), the company is allowed to reduce the fund when it has a loss (deficit) in the current year. My interpretation of the law is that such a reduction prohibits the carry-forward of the same loss for deduction in subsequent years' taxes. The assumption will be that only the non-tax-deductible part of a loss will be deducted from the reserve fund.26

This completes a brief discussion of the company's financial policy, and the various constraints on this policy. Each set of year-end variables has to be determined from systems of simultaneous linear equations. The reasons for this is that both the reserve fund requirements, the state tax, and the capital tax, depend on the other year-end variables, furthermore that different constraints on the balance sheet variables are effective in different cases, and finally that different rules apply to the determination of the reserve fund in different cases. The different constraints give seventeen different systems of linear equations, which are given in Appendix 1. The eight variables that are inter-dependent, endogenous in the systems of equations, are $CSH_t$, $DEB_t$, $KST_t$, ...
RSF\textsubscript{t}, DIV\textsubscript{t}, NDD\textsubscript{t}, the state tax, and the capital tax.

3. The Valuation Model.

In order to evaluate the incentive effects of the tax system, it is necessary to value the tax claim and/or the after-tax cash flows under uncertainty. This will be done once for each of a range of different development plans, to be described in section 4 below. Each plan has non-stochastic vectors of investment costs, operating costs, clean-up costs, and extracted quantities, with one element for each year in the life of the plan. The output prices, however, are uncertain. For simplicity, there will be only one type of output, crude oil, and the royalty will follow the progressive scale.

The\textsuperscript{27} oil price, P, is supposed to follow a lognormal diffusion in continuous time, i.e.

\begin{equation}
\frac{dP}{P} = \alpha P dt + \sigma P dz,
\end{equation}

where \(\alpha\) is the expected instantaneous rate of growth of the price, \(dz\) is the increment to a standard Wiener process, and \(\sigma\) is the instantaneous standard deviation of the price.\textsuperscript{28}

The standard method for valuation under uncertainty relies on some estimate of the systematic risk of the uncertain cash flow, and uses this to obtain a risk-adjusted discount rate or
a risk correction factor for the cash flow. When cash flows are complicated non-linear functions of uncertain prices, this method is hardly applicable, since it is not possible to transform the systematic risk of the prices into the correct systematic risk of the cash flows.

Many tax systems are not in all respects linear in the pre-tax cash flows. A simple example is that positive profits are taxed, whereas losses are not at all refunded. (Perhaps they can be carried forward.) At the time when the uncertainty of the cash flow is resolved, the tax is equal to the maximum of zero and the tax rate times the profit. This is similar to the value at the expiration date of a call option on the tax rate fraction of the profits with striking price zero. The theory for pricing of options, or more generally, contingent claims, can be applied to value such a tax claim.

It should be observed that contingent claims valuation techniques have been applied to the valuation of other aspects of natural resource investments. Even in the absence of taxes, there are many "options" during the history of a license area, such as the option to explore, the option to develop, and the option to extract. These have properties that make the contingent claims analysis necessary, but complicated. Three early contributions in this literature were Tourinho (1979), Paddock, Siegel and Smith (1984), and
Brennan and Schwartz (1985a).

The present paper will not analyze such real investment options. As an approximation, the operation of the oil field is taken as given once the decision to develop has been made, and the valuation in this paper is contingent on such a decision being made. This is in line with some other papers in the recent literature on taxation under uncertainty.

Ball and Bowers (1982) analyze the distortionary effect of the "Resource Rent Tax", a tax system proposed by Garnaut and Clunies Ross (1975). They adapt the Black-Scholes option valuation formula to show that if the tax is applied project by project, the loss offset is quite insufficient, and although the tax is neutral under full certainty, there is a substantial distortionary effect when the probability of a loss is significant.

MacKie-Mason (1986) analyzes the value of the depletion allowance in the U.S. taxation of resource extracting companies using similar methods. The tax code is relatively simple in this case, which allows for a closed-form solution for the value.

Majd and Myers (1985) analyze the value of the tax claim under ordinary U.S. corporate taxation, concentrating on the carry-
back and carry-forward of losses. The rules are so complicated that a closed-form solution does not exist. Instead, they use a simulation technique which is known to converge to the correct value. The technique is Monte Carlo simulation with a very specific transformation of the variables, about which more will be said below.

Jacoby and Laughton (1985) use the same kind of simulation technique to obtain the values of the tax claims on petroleum extracting companies in the U.K. The U.K. petroleum tax code is even more complicated than the Norwegian, and there is no way to get a closed form solution.

The Norwegian system, which is the subject of the present study, also contains non-linearities which are too complicated to allow for closed form solutions. This is true in particular because of the carry-forward of losses and because of the complicated reserve fund requirement.

The method applied by Majd and Myers (1985) and Jacoby and Laughton (1985) circumvents these problems. It originated in work by Cox and Ross (1976), and also follows from Lemma 4 in Cox, Ingersoll and Ross (1985). Its first implementation via Monte Carlo simulation is due to Boyle (1977). Since the method is already known from the literature, I will only give an informal explanation of it.
The method relies on a characteristic of most option valuation methods, starting with Black and Scholes (1973), namely that one can show that the option has the same payoff structure (in all states of the world) as a particular portfolio of a risk-free asset and the underlying uncertain asset(s). The value of the option (today) must be the same as the value of such a portfolio (today). The different underlying uncertain assets in the present study will be claims to units of oil at different specified dates in the future.

An assumption in Black and Scholes (1973) was that the price of the underlying asset followed a diffusion process. By continuous trading it is then possible to adjust the portfolio through time so that the correct payoff is achieved. The adjustments will be self-financing.

The composition of the portfolio does not depend on the expected rate of growth in the prices of the underlying uncertain assets. This opens a possibility for a particular modelling technique. If one in the model simultaneously changes that rate of growth, and makes sure that the market value today of the underlying assets remains the same, the composition of the portfolio, and its value, remain unchanged. This is exactly what the Cox and Ross (1976) method does.
The idea is to transform the uncertain asset prices into processes that have the expected rate of growth that they would have in a world of risk-neutral individuals. 29 The standard assumption is that the assets are such that investors are willing to include them in their portfolios in market equilibrium. Then the "risk-neutral expected rate of growth" of their prices is simply the risk-free rate of interest, which is assumed to be non-stochastic over time, and the starting values for the prices are today's prices. If, on the other hand, investors do not hold the assets in equilibrium, matters are a bit more complicated. I will return to this below.

Even in this transformed world, it will not be possible to derive a closed form solution to what the values of the complicated tax claims will be, since it will be impossible to solve analytically for the composition of the equivalent portfolios. However, it is not necessary in the transformed world to solve for this.

The reason is that the transformed world consists of risk-neutral individuals, and they value cash flows by considering their expected values only, time-discounting at the (only, risk-free) interest rate. The expected values of the complicated non-linear functions of the uncertain prices can be estimated by Monte Carlo simulations. These are also the
values of the equivalent portfolios, both in a risk-neutral world and in a risk-averse world.

This completes the standard description of the method. Crude oil does not fit into the model as an asset that investors are generally willing to hold in equilibrium. Extracted crude oil earns a below-equilibrium expected rate of return. This is similar to the dividend payout of a stock, which affects option value, since the stock without the dividend earns a below-equilibrium expected rate of return.

The most elegant solution to the problem of the expected return shortfall is to assume that there exist futures contracts for all future dates. The futures contracts do earn equilibrium expected rates of return since investors are willing to hold them. Brennan and Schwartz (1985a) give the following relation at time zero between the futures price, \( F \), of a futures contract on e.g. oil, maturing at time \( t \), and the spot price at time zero, \( P_0 \),

\[
(10) \quad F = P_0 e^{(r-y)t},
\]

where \( r \) is the risk-free interest rate, and \( y \) is the convenience yield on oil. This yield can be interpreted as the difference between the expected rate of growth in the oil price, \( \rho \), that would have been sufficient to make investors in
general willing to hold extracted oil in their portfolios, and the actual expected rate of growth of the oil price, $\alpha$. The assumption in this study, as in Brennan and Schwartz (1985a), is that the convenience yield is proportional to the value of oil, with a rate, $y$, which is constant over time.\(^{31}\)

The futures contract, if rolled over to maturity, implies a non-stochastic payment, $F$, in the future. To obtain the value if the payment were due today, discount by the risk-free rate. Thus, the present value of the futures contract, if entered into at time zero, is $P_0 e^{-yt}$. It is easy to show that the expected rate of growth of such a present value, keeping the maturity date fixed, is equal to $\rho$.\(^{32}\) The rate $\rho$ is thus the required expected rate of return on holding oil, or any contract which is perfectly correlated with the return on oil, which is the case with the futures contract under the assumption of a constant convenience yield. Such holding has systematic risk, and $\rho$ is assumed to include a risk premium over the risk-free rate $r$.

In the transformed, risk-neutral world, the present values of futures contracts for each of the maturity dates are the same as in the real world, i.e. $P_0 e^{-yt}$, but the expected rate of growth of these values is $r$, the risk-free rate. For the Monte Carlo simulation, there is in principle a need for one futures contract for each future year, i.e. with maturity
dates at each date when there is a cash flow in the model. The contingent claims valuation of the tax claim can be based on the assumption that the investors are able to hold, and continuously adjust the holding of, all such futures contracts.

However, under the given assumptions it is not necessary in the simulation to keep track of the value of each futures contract separately. Each contract will only affect the calculations at its maturity date, and it is sufficient with a random variable which has the expected value at time $t$ equal to $P_0 e^{(r-y)t}$, and which is perfectly correlated with the return on oil. This can be obtained by simulating the oil price movement, with a price at time zero of $P_0$, and an expected rate of growth of $r-y$ instead of $\alpha$.\footnote{33}

If there exist no contracts or assets which earn equilibrium expected rates of return, and which are perfectly correlated with the return on oil, then contingent claims valuation of the tax claim cannot be based on the argument that investors can replicate the claim by continuously adjusting a portfolio of such assets. McDonald and Siegel (1984) argue that it is still possible to value contingent claims, using a model which characterizes the (required) expected rates of return in a capital market equilibrium.\footnote{34} Such equilibrium models of capital markets are based on much stronger assumptions than
the replication argument, which relies solely on the assumptions of a given stochastic process for the price, on trading being continuous, and on the absence of arbitrage opportunities.

Instead of elaborating on a capital market equilibrium model, I simply assume that there exist assets, traded in capital markets, which make it possible for investors to replicate the returns from holding the futures contracts in question. Then the cash flows to be valued here can also be replicated. Since this study considers U.S. owners, and the tax claim is valued from their points of view, the assumption is that such assets exist in U.S. capital markets. The assets may e.g. be shares in oil companies.

Observe that in the special case when the claim to be valued is linear in the oil price, the valuation from the Monte Carlo simulation has a simple and familiar result. The expected rate of growth of the oil price in the simulations is \( r - y \), and when this is discounted at the rate \( r \), then a unit of oil at date \( t \) in the future has the value at time zero of \( P_0 e^{-yt} = P_0 e^{(\alpha - \rho)t} \). This is equivalent to the standard technique of first estimating an expected rate of growth of the oil price, \( \alpha \), and then discounting the result at the appropriate risk-adjusted discount rate, \( \rho \). I will call this the adjusted present value, APV, of the unit of oil.\footnote{To}
calculate the APV of the before-tax revenues, numbers for \( \rho \) and \( \alpha \) are not needed, only for their difference. 36

A necessary ingredient in the Monte Carlo simulation model is that all decisions that are actually made by firms, can be written down as rules that depend on the outcomes of uncertain variables up to that point in time only. The simulation must describe an event tree, not a decision tree. This shows why it is necessary to describe the company’s financial policy by such rules, as it was done in section 2.

In order for this valuation method to be valid, the expected values of all future years’ tax cash flows, or after-tax cash flows, must at each point in time be a continuous function of the oil price. Only then are investors able to adjust the replicating portfolio in a self-financing way. I assume that the Norwegian system is continuous in this sense, but I have not been able to prove it. 37

From the Monte Carlo simulation, the expected values of the before-tax cash flow, the tax cash flow, and the after-tax cash flow, are estimated for each year in the company’s life. The present values of these are found by discounting at the nominal risk-free rate. 38 In the simulations, the relative oil price increases, \( P_t/P_{t-1} \), are lognormally distributed.
The expected value of \( P_t / P_{t-1} \) is \( e^{r-y} \) for all \( t \), where \( r \) is the nominal risk-free rate.\(^{39}\) The standard deviation of \( \ln(P_t / P_{t-1}) \) is estimated as described in section 5. Values for the logarithm of the increases, \( \ln(P_t / P_{t-1}) \), are generated by a normal random number generator.\(^{40}\)

4. **Data.**

The production possibilities are summarized in alternative development plans, all for the same oil field. In some of the simulations the company is assumed to own shares in more than one field. This is necessary in order to investigate the tax effects of deducting costs at one field from income at another field. The simulations have a base case, and three alternative cases as sensitivity analysis. For each of these four cases, alternative scales of field development are simulated in order to find the after-tax best alternative.

Field development plans are based on the analytic production function

\[
Q = AC^B,
\]

where \( Q \) is a measure of quantity, \( C \) is a measure of costs, \( A \) and \( B \) are parameters, \( A > 0, 0 < B < 1 \). The latter assumption means that there are decreasing returns to scale for alternative development plans for any given field. This is a realistic
description at least from some scale on, while for small-scale
development (especially in deep waters, when fixed costs are
high) there may in fact be increasing returns. This will be
neglected.

The total costs, C, over the life-time of the field are
measured as a present value one year before development start
(time zero). Since factor prices and quantities are assumed
to be non-stochastic, the risk-free interest rate is used as
the discount rate. Since each development plan is totally
non-flexible, there is no need to consider the distinction
between variable and fixed costs. The costs C include
investment (or "development"), operating costs, and clean-up
costs, all before tax.

The total quantity, Q, is measured as an adjusted present
value divided by the oil price at time zero. That is, if \( q_t \)
is the quantity at time t, then Q is \( (1/P_0) \) times the APV of
the cash flow stream \( \{P_t q_t\} \) over the life-time of the field. \( 41 \)
The total APV, or profits, of the field to the company is thus

\[
\Pi = P_0 Q - C = P_0 AC^B - C.
\]

In the absence of taxes, the company would simply maximize
profits by setting the first-order derivative of (12) with
respect to C equal to zero.
The different development plans are obtained by scaling all costs, i.e. all kinds of costs in all years, up or down with the same factor. The quantities $q_t$ are all scaled up or down with another factor, given by (11), the same for all years. This description of the production possibilities represents a compromise between realism on one hand, and generality, availability of data, and simplicity, on the other. In fact, oil companies will seldom consider whether to build (say) three or four identical platforms on a field, especially not under the assumption that each will incur the same time sequence of operating costs in both cases. Even if those were the alternatives, the effect on total output would not be to scale it up or down with the same factor (which is $(3/4)^B$ in the given example) in each year.

However, the actual cost minimizing plans for different levels of $Q$ will have different relations to each other for each individual field, and the resulting time profiles of quantities will also be very field specific. There are no data available, to my knowledge, that allow for a very much more reliable description of production possibilities. The production function (11) is intuitive to economists, and should roughly capture the possibilities for a representative field within the economically relevant range.
The data for investment cost profiles, operating cost profiles, and clean-up costs were calculated as follows from analysis on 33 oil-producing fields in the British and Norwegian sectors of the North Sea (Wood, MacKenzie & Co. (1985)): The investment period is 11 years, and the extraction period is 18 years, with a seven years' overlap, i.e. extraction starts after 4 years of investment. These are average numbers for the 33 fields. Total lifetime of a field is thus 22 years. The time distribution of the investment costs over 11 years, and the operating costs over 18 years, are as follows, representing average time profiles, and normalized to sums of 100: Investment: 7, 16, 16, 16, 13, 11, 7, 5, 4, 3, 2. Operating costs: 3, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 5, 5, 5, 5. Clean-up costs were incurred in the last year of extraction only.

The relation between the three types of costs were determined as arithmetic averages of ratios for the 33 fields, operating costs/investment, and clean-up costs/investment, based on real (constant dollar) values, using a U.S. implicit GNP deflator. The total real value of operating costs was found to be approximately equal, on average, to the total real value of investment, while clean-up costs were found to be 12.5% of investment costs, typically.

The time profile of extraction, normalized to a sum of 100,
was set to 3, 7, 10, 10, 10, 10, 8, 7, 6, 5, 5, 4, 4, 3, 3, 2, 2, 1. This represents a typical build-up, plateau period of production, and exponential decline due to falling pressure.

The value of B is set at 0.55. This is the inverse of the elasticity of a cost function which was estimated in Lund (1984). The estimate has a very large standard error, and it is certainly desirable to do more research on the elasticity of the production function for various types of fields. As a sensitivity test, I have also tried an elasticity of 0.35.

It remains to set the multiplicative constant term, A. That is, in effect, to choose the size of the field. Instead of using the average value in the data in Lund (1984), I have chosen to set A such that in the base case, a before-tax optimal field development reaches the third highest rate on the progressive royalty scale, 12%. Observe that the production function (11) will refer to the company’s 30% share of each field. I have chosen A such that in my base case, the before-tax optimal field development has costs for the 30% share, C = 1.2 billion year-zero U.S. dollars. This optimal development is certainly somewhat larger than the median North Sea field, but not much larger than the mean. (There are some large fields, many small.)
The quantity measure $Q$ does not only depend on the quantities $\{q_t\}$, but also on the discount rate, i.e. on the convenience yield, $y$. The simulations are done for two different values of the convenience yield. In order for the same physical input, or costs, to result in the same physical output, $\{q_t\}$, the total quantity measure, $Q$, is then changed by changing the constant $A$ accordingly. That is, the parameter $A$ in effect becomes a function of the convenience yield $y$.

The real risk-free interest rate is set to 3% per year. This is rather high by historical standards, but not compared to recent rates, and it is not an unreasonable number compared to observed data in the British market for indexed bonds.\(^{47}\)

Both the real risk-free rate and the inflation rate are assumed throughout this study to be certain, and constant. The inflation rate is set at 5% per year.\(^{48}\)

The convenience yield on oil could have been observed if the futures markets had been well enough developed. The longest contracts, however, have maturity dates approximately one year into the future, while our development plans last for 23 years. If the assumption that the yield is constant over time is taken literally, a one-year contract is sufficient to find the yield. The futures market\(^{49}\) for crude oil was studied over 9 months, from November 1985 to August 1986. Only
contracts of approximately one year's maturity were studied, together with the simultaneous yields on T-bills of the same maturity, and the simultaneous spot price of crude oil of the same quality. The convenience yield for one-year contracts varied substantially, from -4% to 30%, with an average over 37 weeks of 13%.

This raises serious doubts about the assumption that the yield is constant over time. The numbers are also surprisingly high. However, there are reasons to believe that the yield would have been closer to constant for contracts of longer maturity, since these would not be as strongly affected by sudden, unexpected shortages or oversupplies in the market. In fact, there may have been strong beliefs in the market during the nine months that the oil price should fall, as it did, and only those with a very high marginal yield from the convenience of storing oil (e.g., governments and oil refineries) would store it in this situation. Also, these would store less than otherwise, to bring their marginal yield up to the(ir) expected rate-of-return shortfall in this period.

By appealing to capital market equilibrium models, such as the intertemporal versions of the CAPM, it is quite reasonable to believe that the required expected rate of return on oil, \( \rho \), is constant over time in the longer run (if oil's
systematic risk and the market price of risk are both constant). The actual expected rate of return (or growth) of the oil price, $\alpha$, may vary a lot in the short run, but is not likely to vary much in the long run, unless one has very specific views on periodical movements in the cartelization of the market. 51

I conclude that the assumption of a constant convenience yield has some arguments for it in the longer run. Also, I do not know any good alternative. It remains, however, to find numbers to use in the simulations. I choose 4% as my base case. This is the average over six weeks from the beginning of July 1986, when the oil price had stabilized. It is also approximately equal to the convenience yield in the U.S. market for developed oil reserves, according to Paddock, Siegel and Smith (1984). 52 In order to test the sensitivity of my results, I have used 6% as an alternative.

The standard deviation of $\ln(P_t/P_{t-1})$ is estimated from monthly data over the period July 1973 - March 1986, as 0.37 on a yearly basis. 53 The crude oil price in year zero of the model, which is the actual year 1986, is taken to be 15 U.S. dollars per barrel.

For this $P_0$ of 15 dollars, a 30% share of the the before-tax
optimal development of the field to be studied has a cost $C$ of 1.2 billion year-zero dollars.

The company is assumed to own 30% shares in the fields. This is in line with the Norwegian practice of giving licenses to groups of companies. A typical share to the largest foreign participant in later years has been 30%. The company is assumed to have a debt/equity maximum, DEM, of 3, i.e. a debt maximum of 75% of assets. This is the assumption in the official Norwegian evaluations of the tax system.$^{54}$

The valuations were done under four sets of assumptions, called cases I - IV. In case I, the company has shares in two other fields, and the development of these started ten and five years, respectively, before the field under consideration. The development plans for the two other fields are kept fixed. The purpose of including those fields, is that in this way, the company for many oil price sequences will be in tax position when the new field has negative cash flows. This makes the carry-forward of losses and uplift unnecessary, and the after-tax value of the new field increases. Furthermore, the company will have a reserve fund of some size when it considers the new field, and this reduces its effective cost of capital. The choice situation to be analyzed is what scale of development of the third field the company would want, given that it already has shares in the
first two. The two other fields are both of the case I before-tax optimal size, i.e. with costs of 1.2 billion year-zero dollars, and the resulting case I physical output. This choice is rather arbitrary. Moreover, case I has a convenience yield of 4% and an elasticity of the production function of 0.55.

Case II is like case I, but with the company having a share in only one field. Case III is like case I, but with a convenience yield of 6%. Case IV is like case I, but with an elasticity of the production function of 0.35. With this elasticity, the constant \( A \) is adjusted such that the development scale of 1.2 billion year-zero dollars still gives the same physical output. That is, this point, as well as \((0,0)\), is still on the production function. Smaller scales of development will thus give a higher output than with the case I production function, and larger scales will give a relatively lower output.

The before-tax optimal development plan has a cost of 1.2 billion year-zero dollars in cases I and II. In case III, the before-tax optimal development plan is approximately 0.75 billion year-zero dollars, and in case IV, it is approximately 0.6 billion. In all four cases, I have looked for the after-tax optimal scale of development by doing a grid search, where the grid size has been 100 million dollars. That is, a
development plan costing 100 million dollars has been simulated, next a plan costing 200 million dollars, and so on.

5. Results.
The computer program has been running on computers of the IBM PC type, and it has turned out to be very time-consuming. Each run of 22 years takes about 10 seconds on an IBM PC/AT, and in order to get sufficient precision, one sometimes needs more than ten thousand runs. A run of three fields takes 14 seconds. The standard simulation length has been ten thousand runs, taking about 20 hours on two computers. 58

The central result from each simulation is the estimate for the tax value, or for the after-tax value to the company, of the specified development of the field. The simulation also gives an estimate for the before-tax value, but this value is known exactly, since it is an APV. This provides a good check on whether the results are reasonable. In all simulations, the estimate of the APV has been closer to the true APV than the standard error of the estimate.

Since the estimate is the sum of a large number of drawings of independent, identically distributed random variables, it is asymptotically normally distributed. It is then possible to estimate 95% confidence intervals for the estimated values. In order to conclude that the company would have preferred one
development plan to some other under some set of assumptions, I want the estimated 95% confidence intervals of the plans' estimated after-tax values to be non-overlapping. 59

The estimates for the after-tax values in case I are given in table 1, together with their estimated 95% confidence intervals, and the APV of the before-tax values. The numbers are shown for five different scales of development.

<table>
<thead>
<tr>
<th>Cost of development</th>
<th>Before-tax APV</th>
<th>After-tax value</th>
<th>95% conf. interval of after-tax value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1200</td>
<td>982</td>
<td>-224</td>
<td>( -266, -182)</td>
</tr>
<tr>
<td>600</td>
<td>890</td>
<td>8</td>
<td>( -15, 31)</td>
</tr>
<tr>
<td>300 *)</td>
<td>718</td>
<td>85</td>
<td>( 69, 101)</td>
</tr>
<tr>
<td>200 *)</td>
<td>614</td>
<td>99</td>
<td>( 79, 120)</td>
</tr>
<tr>
<td>100</td>
<td>456</td>
<td>98</td>
<td>( 81, 115)</td>
</tr>
</tbody>
</table>

Table 1: Results from simulations in case I. (Other activity, y=0.04, B=0.55.) Numbers in million 1986 dollars.

*) Simulation consists of 15000 runs.
The results show very clearly that the tax system has distortionary effects under these assumptions. Even though the 1.2 billion scale is before-tax optimal, it has a negative after-tax value. The company has an incentive to choose a development plan which is less than half as big. However, the results do not bring out a clear answer to which development scale is after-tax optimal. The confidence intervals for the values of the 100 million, 200 million, and 300 million scales overlap each other. In particular, it is difficult to answer the question of which of the 100 million and 200 million scales is the better.

Even to get a clear answer to which is the better of 200 and 300 million, the computer program would have to run for approximately one month continuously on an IBM AT with a 4MHz processor. 60

One preliminary conclusion is that the method or the speed of computation needs to be improved in order to find more precise estimates.

Table 2 shows the corresponding results in case II, i.e. the same assumptions as in case I, except that the company has no other activity. A priori this is believed to reduce the after-tax value of the field to the company, since the tax deductions for costs will be effective at later dates for
almost all realized paths of the oil price.

<table>
<thead>
<tr>
<th>Cost of development</th>
<th>Before-tax APV</th>
<th>After-tax value</th>
<th>95% conf. interval of after-tax value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1200</td>
<td>982</td>
<td>-249</td>
<td>(-265, -233)</td>
</tr>
<tr>
<td>600</td>
<td>890</td>
<td>-10</td>
<td>(-19, -1)</td>
</tr>
<tr>
<td>300 *)</td>
<td>718</td>
<td>61</td>
<td>(57, 65)</td>
</tr>
<tr>
<td>200 *)</td>
<td>614</td>
<td>69</td>
<td>(66, 72)</td>
</tr>
<tr>
<td>100</td>
<td>456</td>
<td>63</td>
<td>(60, 66)</td>
</tr>
</tbody>
</table>

Table 2: Results from simulations in case II.  
(No other activity, y=0.04, B=0.55.)  
Numbers in million 1986 dollars.  
*) Simulation consists of 17000 runs.

Table 2 shows that the after-tax values at all scales of development are less than the after-tax values of the corresponding scales in table 1. This confirms the a priori belief that a company which has no other activity will value the field lower than a company with shares in other activity within this tax jurisdiction. There are strong distortionary effects, as in table 1, and the effects are very similar.
That is, with or without other activity, the scale which gives the highest estimated value is 200 million dollars, or one sixth of the before-tax optimal field. One may say that the estimated relative distortion is one sixth. The reduction in after-tax values compared to table 1 is between 14 and 35 million dollars at all the scales in the table, with no monotonic variation according to scale.

Looking at the confidence intervals, table 2 shows that by increasing the number of runs to 17000, it was possible to obtain non-overlapping confidence intervals between the scale of 200 million and the two neighboring scales in the grid search. Of the simulated scales, I can conclude at a 95% confidence level that the scale of 200 million dollars has the highest after-tax value in case II.

Table 3 shows the results in case III. This is equivalent to case I, except that the convenience yield on oil has been increased to 0.06. The production function is the same in the physical sense, but since Q is discounted by the convenience yield, the multiplicative constant in the production function, A, has been reduced correspondingly.

Table 3 shows that the highest after-tax value with this convenience yield is obtained at a scale of 100 million, and that the confidence interval of this value almost does not
<table>
<thead>
<tr>
<th>Cost of development</th>
<th>Before-tax APV</th>
<th>After-tax value</th>
<th>95% conf. interval of after-tax value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1200</td>
<td>567</td>
<td>-355</td>
<td>(-383, -328)</td>
</tr>
<tr>
<td>600</td>
<td>607</td>
<td>-58</td>
<td>(-82, -33)</td>
</tr>
<tr>
<td>300</td>
<td>524</td>
<td>56</td>
<td>(38, 74)</td>
</tr>
<tr>
<td>200</td>
<td>459</td>
<td>79</td>
<td>(62, 96)</td>
</tr>
<tr>
<td>100</td>
<td>350</td>
<td>89</td>
<td>(73, 105)</td>
</tr>
</tbody>
</table>

Table 3: Results from simulations in case III. 
(Other activity, y=0.06, B=0.55.)
Numbers in million 1986 dollars.

overlap the confidence of the after-tax value of a 300 million dollars development. This suggests that the after-tax optimal scale is less than 300 million dollars in this case. If the estimates in table 3 are compared to those in table 1, neglecting for a moment their standard errors, it seems that the after-tax optimal scale is somewhat smaller in case III than in case I. But so is the before-tax optimal scale, 750 million dollars compared to 1200 million dollars. A conclusion is that the relative distortions in the two cases seems to be approximately the same. Considering the standard
errors of the estimates, it is at least clear that such a conclusion cannot be rejected.

<table>
<thead>
<tr>
<th>Cost of development</th>
<th>Before-tax APV</th>
<th>After-tax value</th>
<th>95% conf. interval of after-tax value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1200</td>
<td>982</td>
<td>-224</td>
<td>( -266, -182)</td>
</tr>
<tr>
<td>600</td>
<td>1112</td>
<td>66</td>
<td>( 43, 90)</td>
</tr>
<tr>
<td>300</td>
<td>1043</td>
<td>166</td>
<td>( 146, 187)</td>
</tr>
<tr>
<td>200</td>
<td>965</td>
<td>184</td>
<td>( 165, 203)</td>
</tr>
<tr>
<td>100</td>
<td>814</td>
<td>189</td>
<td>( 171, 207)</td>
</tr>
</tbody>
</table>

Table 4: Results from simulations in case IV. 
(Other activity, y=0.04, B=0.35.) 
Numbers in million 1986 dollars.

Table 4 shows the results in case IV. This is equivalent to case I, except that the elasticity, B, of the production function has been reduced from 0.55 to 0.35. Also, the constant term, A, in the production function has been changed so that the development scale of 1200 million dollars still gives the same output.
Table 4 shows, like table 3, that the highest after-tax value is obtained at a scale of 100 million dollars, but again, the 95% confidence interval of this estimate overlaps the intervals of other estimates. The before-tax optimal development in this case is about 600 million dollars, and I cannot reject an hypothesis that the relative distortion is the same as in case I.

Figure 1 shows a comparison of cases I and II. They have the same before-tax values, but the after-tax value is higher for case I at all scales above zero. Figure 1, as well as figures 2 and 3, does not show the standard errors or confidence intervals of the estimates, only the estimated values. The curves that are drawn through the after-tax values are of course tentative, in the sense that I have not investigated their shapes outside the five points that are shown. There is, however, every reason to believe that the curves should be close to continuous, apart from the small discontinuity in the royalty scale.

Figure 2 shows a comparison of cases I and III. The characteristics of the before-tax curves are recognized in the after-tax curves. That is, the sensitivity analysis has not given any surprises
Figure 1: Before-and-after-tax values in cases I and II
(y = 0.04, B = 0.55)
——: Before-tax value.
——: After-tax value with other activity.
———: After-tax value of stand-alone field.
(Numbers in billion 1986 dollars.)
Figure 2: Before - and after - tax values in cases I og III
(other activity, B = 0.55).

---: Convenience yield y = 0.04.
-----: Convenience yield y = 0.06.

(Numbers in billion 1986 dollars.)
Figure 3: Before – and after – tax values in cases I and IV
(other activity, \( y = 0.04 \)).

---: Elasticity B = 0.55.

----: Elasticity B = 0.35.

(Numbers in billion 1986 dollars.)
Figure 3 shows a comparison of cases I and IV. The figure illustrates the way in which the before-tax value as a function of costs has been changed. The elasticity has been lowered, while the 1.2 billion development still gives the same before-tax value. Again, the after-tax value curves relate to the before-tax value curves in non-surprising ways.

6. Discussion.

The contingent claims valuation method has shown that Norwegian petroleum taxation has severe distortionary effects under the given assumptions. These effects seem to be stronger than those found by Asheim and Furset (1982), although no direct comparison is possible, since Asheim and Furset described the production possibilities for the field differently. It is likely that stronger effects will be found by this study, taking uncertainty into account, since the possibility of never earning tax deductions is seldom well accounted for in studies that assume full certainty.

Due to limitations on the number of simulation runs, the distortionary effects cannot be stated very precisely. For the given assumptions, the best estimate is that the companies will want development plans of about one sixth the size of a before-tax optimal plan. At a 95% confidence level, the relative distortion is at least one half in the cases with an elasticity of the production function of 0.55. In table 4,
the case with an elasticity of 0.35, the precision of the estimates is less, so that even though the best estimate of the after-tax optimal development, 100 million dollars, is one sixth of the before-tax optimal development, 600 million dollars, I cannot conclude on a 95% confidence level that the 100 million development is better after tax than the 300 million development scale.

The term relative distortion has been used as the after-tax optimal development costs divided by the before-tax optimal development costs. However, a more meaningful number may be the after-tax optimal before-tax value divided by the before-tax optimal before-tax value. After all, the costs are no goal in themselves, whereas the before-tax value may be taken as the total value to the (world) society of the field.

The relative value distortions, based on the developments that have been found after-tax optimal in the grid search, and the known before-tax APVs of these, are 614/982=0.63 in cases I and II, 350/614=0.57 in case III (in which 750 million dollars development gives a 614 million dollars APV), and 814/1112=0.73 in case IV.

A cash-flow tax is the only system that is known under uncertainty to leave the after-tax optimal development equal to the before-tax optimal development. A revenue-neutral
reform to a cash-flow tax would require tax rates in the four cases of \((614-99)/982 = 52.4\%\) (case I), \((614-69)/982 = 55.5\%\) (case II), \((350-89)/982 = 26.6\%\) (case III), \((814-189)/1112 = 56.2\%\) (case IV). Because of the low precision of the estimated after-tax values, these numbers should be interpreted with care. One conclusion is obvious, however, namely that a revenue-neutral reform will look differently in the four cases.

It is also possible to calculate a profits-neutral reform, i.e. rates of a cash-flow tax that would leave the companies as well off as under the present tax regime. These rates are

\[
\begin{align*}
1-99/982 &= 89.9\% \text{ (case I)}, \\
1-69/982 &= 92.9\% \text{ (case II)}, \\
1-89/614 &= 85.5\% \text{ (case III)}, \\
1-189/1112 &= 83.0\% \text{ (case IV)}. 
\end{align*}
\]

If more cases had been analyzed, if the grid in the grid search had been denser, and if more precise estimates had been obtained, the tax reform numbers would have been more interesting. It is likely that both the revenue-neutral reform rates and the profits-neutral reform rates would have varied a lot. However, if they covered disjunct intervals (as above, where the intervals are 26.6\% to 56.2\% and 83.0\% to 92.9\%), one could design a Pareto-improving reform by choosing a cash flow tax rate between the intervals. In all the cases that have been studied here, any cash flow tax rate between 56.2\% and 83.0\% would have left both the state and the
companies better off. An interesting question is whether the intervals would in fact be disjunct if a wide range of cases had been covered.

A natural extension of the study would be to test the sensitivity of the results to the standard deviation of the oil price. An easy case to start with would be to reduce it to zero, in order to bring out the difference between certainty and uncertainty for this analysis.

Furthermore, it would be interesting to test the different elements of the tax system. The present study has not indicated which elements have the strongest effects in any sense. For a gradual reform of the tax system, this would clearly be of importance.

For a discussion of a tax reform, one must also consider the other groups of decisions that were mentioned in section 1.

The different assumptions that have been made are all open to criticism. One weak point is the estimate of the elasticity of the production function. But it is reassuring that the relative distortion seems to be the same with both the two elasticities that were used.

In order to interpret the results, they should be looked at as
describing the companies' incentives, not as an hypothesis about what has actually happened in the Norwegian petroleum sector. Because of the state participation, the actual decisions regarding development will be compromises between Statoil and other participants.

A seemingly counter-acting incentive effect of the tax system is that its high rates encourages foreign companies to incur costs of technological experiments and of training. If the benefits of these can be transferred to countries or sectors with lower tax rates, the companies will be interested in using the Norwegian sector as a "playground". The model of this study does not account for that, but it is discussed in Lund (1985). Of course, the costs incurred because of this motive will not lead to solutions that are before-tax optimal.

One problem in the study has been to check whether the results were in fact correct. There are two partial checks of this. One is that the computer program gives tax values and after-tax values that sum up to the before-tax value. It does. The other is that the computer program estimates the before-tax value close to what it is known to be before the simulations. It does. However, there is no test for whether the taxes are actually calculated correctly. I have checked the accounting in all the thirteen cases of simultaneous equations, and it is correct. But there may still be sources of error. This is
one motivation for reproducing the computer program in appendix 2. The reader may check the calculations.

An obvious requirement for extending the study is to speed up the simulations. Unfortunately, I know of no obvious way to write the computer program more efficiently for this tax system. However, faster computers exist.

7. Conclusion.

Contingent claims valuation techniques have been used to analyze the Norwegian petroleum tax system (as of 1985) under output price uncertainty. Various development plans for the same field have been considered, and the tax system's possible effects on the companies' choice between these plans have been investigated.

The tax system seems to have strong distortionary effects, giving the companies incentives to reduce the development of each field to less than one half the before-tax optimal development size, and before-tax value by 30 to 40 percent.

The method is promising, but computationally very time-consuming. The study has only covered a small range of possible situations, and the results should therefore be interpreted with care.
Notes.

1. E.g., companies decide on how many platforms to install on an offshore oil field in deep water. This decision is not at all trivial.

2. More will be said on exploration incentives below.

3. It may be optimal to start development at a small scale at one point in time, and expand by installing more capacity later. This will be neglected in the present study.

4. Newbery (1986) is quoted as follows: "The key question therefore is not how do you adjust the depletion rate in response to changes in the world price of energy, because by and large once you've put the infrastructure in you get it out as fast as you possibly can. It would take an enormous change in expectations, which in some sense would not be rational at the global level, for you to want to change that." This "fact" depends on technology, which may change as oil prices presently (1986) have fallen sharply, but my model does not account for that.

5. In general, the decisions on exploration and on time for development may be more interesting, see e.g. Newbery (1986).

6. For royalty and taxation purposes the price of oil is set equal to a "Norm Price", determined by a publicly appointed committee, the Petroleum Price Board. The intention is to avoid transfer pricing between companies with common interests. The law requires that the price reflect what the
oil could have been sold for in the market, but the oil companies claim it is a bit too high. This claim will be neglected.

7. The best source in English for a description of the system is probably Syversen (1986).

8. The rate increases with monthly production, and falls accordingly, except that if it has reached 12%, it cannot later fall below 12% for the same field. In this study, the scale will be translated into yearly production for simplicity.

9. In order to get the correct royalty numbers, I will distinguish between production from different fields. Also, for oil subject to progressive royalty, I will adjust the scale according to the company's share of the field.

10. Actually every company is required to have an onshore office, and it will pay some taxes according to onshore rules. But for a company with only a small staff onshore in Norway, these taxes are negligible in comparison to the offshore taxation if petroleum is extracted. Moreover, all activities by foreign companies can for taxation purposes be assumed to be organized as Norwegian corporations, since they are taxed this way, even though they may legally be branches of foreign corporations. (Artificial dividends, reserve funds (see below), etc., are calculated by such branches for taxation purposes.)

11. Details on the rates, the tax base, and deductions for
each tax, are as follows:

1) The royalty. The base is gross revenues after deduction of some transportation costs.

2) The income tax. The base is (yearly, accounting) profits. Deductions are: previous years' losses (at least 15 years carry-forward, can be extended), royalty, operating expenses, net interest expenses, and depreciation, which is six years straight line, of book value, from the time an investment is taken in normal use and petroleum is extracted.

3) The state tax. The base is as for 2), except that dividends paid at the year-end out of the year's profit are deductible. (The background for this is that 2) and 3) constitute the regular taxation of Norwegian onshore companies, 2) being a municipal tax, and 3) being a state tax. Dividends are taxed at a progressive state tax at the personal level. Norwegian personal taxation will not be considered in this study. For foreign companies, tax 5) is supposed to make up for part of the dividend deduction, and replace the taxation of dividends at the personal level. For the offshore petroleum sector, the "municipal" tax is paid to the state.)

4) The special petroleum tax. The base is as for 2), except that an extra depreciation allowance, called the uplift, is deductible. This is determined as for the
standard allowance, except that it is straight line over 15 years, starting the year after the first year of depreciation allowances. To the extent that this allowance cannot be effectively deducted because of low profits in any year, it may be carried forward. This is done independently of the carrying forward of losses.

5) The withholding tax, also called the source tax, is a tax on dividends distributed to foreign shareholders. It is in most cases determined in tax treaties between the two involved nations, and is most often 15%.

6) A capital tax (or wealth tax) is levied on the book value of equity, i.e. assets minus debt. The book value of physical capital is written down according to the depreciation mentioned under 2).

12. In some licenses there is a "sliding scale" participation, according to which Statoil's share will increase if a field is discovered and developed. The final share depends on the maximum yearly extraction rate over the field's life. In this case Statoil shares development and operating costs, but not exploration costs, according to the final participation share. In addition to the distortionary effect of the carried interest, the "sliding scale" has a possible distortionary effect in giving the companies other than Statoil an incentive to reduce the maximum yearly extraction rate.
13. This is a pure book-keeping transaction on the "liabilities" side of the balance sheet, and does not imply the transfer of assets to separate accounts.

14. The regulations also mention a "revaluation fund", which will be neglected here.

15. Norwegian Ministry of Finance (1980, p. 22) writes, "In many cases so-called underfinancing has been claimed against the companies. In some cases - where the lack of equity has been too obvious - the companies have agreed to increase equity. This has been done in many ways, e.g. by issuing shares. It has also happened that the tax assessment authorities have refused the deduction of debt on this background," my translation. "Deduction of debt" should be interpreted as deduction of interest payments.

16. Another result of the generous dividend deduction is that companies prefer to pay very high interest rates if the lender and the borrower have common interests. This is prohibited by the tax authorities, and I shall assume that this regulation effectively prohibits higher interest rates than market rates.

17. There are some minor points that have been neglected. 
1. The grace period for tax payments is on average one half year, except for royalties, for which it is 75 days. In the numerical calculations, I shall neglect the grace period.
2. The tax authorities do not accept interest deductions for companies that have no share in any commercial discovery. The
ruling has been that such companies cannot be debt financed at all on an arms-length basis. 3. Clean-up costs when fields are abandoned are partly refunded by the treasury at a rate equal to the average tax rate over the life-time of the equipment to be removed. This is not part of the ordinary tax scheme. The assumption in this study is that this average rate is \( t_y^t + t_s^t + t_x^t \). 4. For the numerical calculations, I assume that a company is liquidated in the year in which it produces its last quantity of oil. If the uplift allowances that have been effectively deducted do not sum to the total investment costs of the company up to that year, the company is allowed a carry-back of the uplift allowances, getting previous years' special tax payments refunded. Apart from the clean-up cost refund, this is the only case in which there may be negative taxes. I have interpreted the tax law such that the carry-back cannot extend to years before the year in which the respective investment costs have been incurred. (This case has never yet occurred in practice, and the law is not clear.)

Observe that point 2 and 3 imply that there are two kinds of non-deductible expenditures, non-deductible interests, \( NDI_t \), and clean-up costs after refund, \( CCA_t \). The after-tax profits in (5) are equal to \( \text{PBT}_t^t \), as defined in (1), minus the \( t_y^t, t_s^t, t_x^t \), and \( t_k^t \) taxes, minus also \( NDI_t \) and \( CCA_t \).

18. When I discuss the "tax system" in this section, state participation is neglected, but the taxes, the royalty, and the reserve fund requirement are considered unless explicitly
neglected.


20. Exploration costs may be capitalized or expensed. This has little significance. I assume the latter. This means that the exploring company has losses, which are carried forward.

21. In the U.S., foreign tax credits prevent the company from paying any U.S. taxes on profits from its Norwegian subsidiary.

22. The financial assets, $CSH_t$, are assumed to carry interest at the same rate as the debt, $DEB_t$. Both variables are assumed to be non-negative. The reason why they cannot be collapsed into one variable, allowed to take positive or negative values, is the constraint (7).

23. There are legal constraints on the reduction of $KST_t$, but with the consent of the creditors, the capital stock can be reduced at will with a time lag. This time lag will be neglected.

24. There is a less clear-cut case when the company is allowed to reduce its reserve fund because it exceeds the minimum given by (4), and this reduction allows for a dividend distribution exceeding the year's profit. The assumption will be that the company in this case distributes non-deductible (for state tax purposes) dividends, $NDD_t$. 
25. Investment in new real assets may be particularly attractive in this case, with a low effective after-tax cost of capital. The possibility that there are investments in new real assets is partly investigated in this study when the company is assumed to own shares in more than one field.

26. The non-tax-deductible part consists of taxes that are paid even when there is a loss, in particular the capital tax, and of the other non-deductible expenses, $NDI_t$ and $CCA_t$. If the reserve fund is too small to cover such a non-tax-deductible loss, the excess part of the loss might have been accounted for as a negative number on the liabilities side of the balance sheet, just like the tax-deductible loss, $LOS_t$. But I assume an alternative accounting procedure, that $KST_t$ is reduced by this number. This latter assumption has no economic significance.

27. The symbol $P$ in this section has the same meaning as the symbol $PRI$ in the previous section.

28. This assumption is unnecessarily restrictive for the contingent claims valuation technique. It would have been sufficient that the expectations of the prices at all future moments follow diffusion processes, as in Jacoby and Laughton (1985). However, for a numerical application it is difficult to come up with a more realistic specification. The specification here is similar to the one in Brennan and Schwartz (1985a). It corresponds to one of the two alternative price processes in Lund (1986).
29. The assumption for such a transformation is that the returns' (or growth rates') higher-order moments, and comoments with other random variables in the economy, are unchanged. This is also the case in the discussion of the required expected rate of return on oil, $\rho$, below.

30. This is known from the economic theory of natural resources under full certainty. Correspondingly, unextracted crude oil with high extraction costs earns an above-equilibrium expected rate of return, in the sense that it is profitable to postpone extraction.

31. It is possible to define $y$ as $\rho - \alpha$. Alternatively, when futures contracts exist, equation (10) can be taken as the definition of the convenience yield. (If $F$ is not corrected for expected inflation, then $r$ is a nominal rate.) Only investors who actually value the marginal yield from the convenience of storing oil as higher than $y$, will hold, i.e. store, oil. For other investors, there is an expected-rate-of-return shortfall on holding oil, compared to the expected rate of return they would require to hold oil in their portfolios. The shortfall is $\rho - \alpha$ for an investor who experiences no yield at all from the convenience of storing oil. This is similar to the expected-rate-of-return shortfall on holding a dividend-paying stock without receiving the dividends.

32. The present value of a contract at time $1$ with maturity date $t$ is $P_1 e^{-y(t-1)}$. The expected value of this as
of time zero is $P_0e^{\alpha e^{-y(t-1)}} = P_0e^{\alpha+y e^{-yt}}$, and $\alpha+y=r$.

33. Cox, Ingersoll, and Ross (1985) interpret their Lemma 4 as follows, "the equilibrium price of a claim is given by its expected discounted value, with discounting done at the risk free rate, when the expectation is taken with respect to a risk-adjusted process for wealth and the state variables. The risk adjustment is accomplished by reducing the drift of each underlying variable by the corresponding factor risk premium." In this study, the risk premium is $\rho-r=\alpha-(r-y)$, and since the actual expected rate of growth of the oil price is $\alpha$, the drift of the oil price is transformed to $r-y$.

34. Such a model may e.g. be the intertemporal CAPM of Merton (1973). If the difference between the required expected rate of return on oil, $\rho$, and the actual expected rate of return, $\alpha$, is constant, this difference is analogous to a constant convenience yield, and the two alternative approaches yield the same answers for closed-form option values.

35. Lessard (1980) uses the term APV in a broader sense, more or less denoting the result of any "valuation by components". My use of the term is a special case, where the components are distinguished by their systematic risks only. In this study there are two such components of the before-tax cash flow, the revenues and the costs, the latter having no risk. Brealey and Myers (1981) use the term APV somewhat differently.
36. Brennan and Schwartz (1985b): "This approach avoids simultaneously the twin problems of assessing the expected future spot price and of assigning a discount rate appropriate to the risk of these revenues." As long as futures contracts of the necessary maturities exist, one can find the convenience yield from futures market data.

37. The only obvious discontinuity is that the progressive royalty is a discontinuous function of output. This does not represent a problem under my assumptions, since I assume that output is non-stochastic. The other regulations are typically non-linear, but continuous. E.g., as the expected oil price increases, the expected taxable profits may increase from negative to positive. As soon as they become positive, 10% of profits will be channeled into the reserve fund. Furthermore, as soon as they allow for dividends which exceed 10% of the preceding year's (KST+RSF), then an amount equal to the excessive dividends will be channeled into the reserve fund. This leaves expected dividends as a continuous function of the expected oil price.

38. The three present values have the relation that before-tax value = tax value + after-tax value. This follows since the company is assumed to pay (or earn) interest at the risk-free rate on its financial assets, discounting is done at this rate, and the creditors and debtors are the only recipients of cash flows apart from the state and the owners. A good test on the correctness of the calculations is whether
tax and after-tax values actually sum up to before-tax value. This holds for the program in Appendix 2, when the tax value is calculated from the six taxes plus the tax share of the clean-up costs, while the after-tax value is calculated as programmed, i.e. from the dividends and the additions to and subtractions from the capital stock.

39. This means that the expected value of $\ln(P_t/P_{t-1})$ is $r-y-(\text{half the variance of } \ln(P_t/P_{t-1}))$, see e.g. Aitchison and Brown (1957).

40. The random number generator is part of the computer language GAUSS. Standard normal deviates are obtained from uniform deviates by an algorithm proposed by Kinderman and Ramage (1976). The uniform random number generator is based on a method which is found in Kennedy and Gentle (1980).

41. This implies that $Q$ can be seen as the APV of the quantities $\{q_t\}$, discounted at the rate $y=\rho-\alpha$.

42. The elasticity of the cost function was estimated as 1.867, with a standard error of 1.045.

43. Nystad (1985) argues that fields with different geological characteristics have quite different production functions.

44. Unofficial sources have claimed that increasing the number of platforms on the Statfjord field, the largest oil field in Norway, from 2 to 3, increased the planned total output by about 15%. This gives an elasticity of about 0.35. No one has been willing to confirm the number publicly.
45. Total costs for the field is $(1.2/0.3)$ billion dollars.

46. Data for the development plans can be constructed from the given assumptions. Approximate data for the development scale of 1.2 billion dollars (for the company's 30% share), and an inflation rate of 5% per year, are given as an illustration:

Investment: (In million dollars, nominal, starting in year one.) 55, 132, 138, 146, 124, 111, 74, 56, 47, 37, 26.

Operating costs: (In million dollars, nominal, starting in year five.) 29, 50, 63, 67, 70, 74, 77, 81, 86, 90, 95, 99, 105, 110, 96, 101, 106, 112.


Clean-up costs: 12.5% of the sum of investment costs in real terms, incurred in year 22.

47. Blanchard and Summers (1984) refer to an implicit ten-years rate (using data from the indexed bond market) of 3% in 1996, while implicit rates at earlier dates are higher.

48. This is rather high by the most recent U.S. standards, but not very high by Norwegian standards. In fact, a satisfactory treatment of inflation is far beyond the scope of this study. A completely realistic treatment should have taken into account that tax calculations are done in Norwegian
kroner, such that the kroner/dollar exchange rate is important. Also, both this rate and the inflation rate are in fact risky.

49. The market is the New York Mercantile Exchange, and the data source is the Wall Street Journal.


51. One can start the argument referring to exhaustible resource economic equilibrium models under full certainty, with a constant market regime, monopoly or competition. These give a rate-of-return shortfall which depends on the marginal extraction costs. These costs are assumed to increase over time, and can hardly be fluctuating very much. Under uncertainty, high expected fluctuations in the oil price can more easily be ruled out in the longer run, since there are long adjustment lags in the development and extraction process for crude oil. If a resource owner expects an unusually high or low growth in the oil price starting five years and ending seven years from today, there should be plenty of time for adjusting the extraction to take advantage of this.

52. Since developed oil reserves still have a marginal extraction (operating) cost, there is no direct equivalence between the convenience yield on such reserves and on extracted crude oil.

53. Oil prices are nominal prices, spot prices from 1976, for Saudi Arabia Light, as given in Energy Economic Research (1973-1985). The choice of estimation period is not obvious.
It is based on the belief that the period after OPEC was able to cartelize the market for the first time is more representative for the future than the previous period.


55. In order to analyze the value of the new field, the simulations are first done without the new field. Then the different development plans for the new field are combined with the two other fields, and the after-tax value of the new field (for each development plan) is taken to be the difference between the values with and without the new field. For comparability, the oil price will be 15 dollars per barrel in the year before development start of the third field, and the Monte Carlo simulations start from there. Furthermore, the present values are values as of that year. The first ten years of the company's life, when it only participates in the two other fields, are supposed to be 1976-85, and historical oil prices are used for these years.

56. In fact, the companies would probably have chosen to develop its first two fields at scales closer to the after-tax optimum. On the other hand, if these fields were developed starting in 1976 and 1981, the oil prices were higher. The important point is the inclusion of some other activity, while the specifics of that activity in reality will differ a lot between companies. Furthermore, I believe it is more realistic to assume the existence of two fields, which may in some cases result in a company out of tax position, than to
assume that the company is always in tax position, as some studies have done. For very low oil prices, no oil company can be certain that it will be in tax position.

57. In cases I and II, A is 0.0659444 when the units are thousand dollars and million barrels. In case III, A is 0.0533936. In case IV, A is 1.0839615.

58. The random number generator was seeded equally for each simulation of ten thousand runs. Thus, there exists an exact frequency distribution for which the estimates from different simulations can be compared directly, without considering their standard errors. Observe, however, that this is a frequency distribution over sequences of 22 price increases.

59. When the company participates in three fields, it is sufficient to look at the simulations including all three fields, and compare such simulations for different development plans for the third field, to decide on which development plan the company would have chosen for the third field, given that the first two were already developed. In order to get the additional value of the third field, however, the results of these simulations were compared with a simulation without the third field. The estimated value of the two first fields was simply subtracted from the estimated value of the three fields together. For the estimate of this difference, no standard error or confidence interval was estimated.

60. The estimated time of computation is based on the
estimated difference between the after-tax values, the
estimated confidence intervals, and the fact that the width of
confidence interval is proportional to the square root of the
number of runs. From the given numbers, an estimated 90000
runs are needed to get non-overlapping confidence intervals.
(The estimate is not precise, of course. It is based on the
estimated difference between the two after-tax values, but if
this had been known precisely, no more simulations had been
needed.) The estimates for the 100 million and 200 million
scales are so close that it makes little sense to estimate the
necessary time of computation.

6l. That is, each year's extracted quantity, \( q_t \), is the
same as under the first set of assumptions, for any given
scale of development.
References.


Appendix 1:

Details of Accounting and Tax Calculations.

This appendix describes in detail the model of the company's accounting and tax calculations. For these calculations, the "real" variables, i.e. investment, operating, and abandonment costs, as well as the extracted quantities and their prices, are given. The model will for each year determine the company's taxes, its income statement, and its balance sheet. The reason for keeping track of the accounting variables and identities is that these, in particular the reserve fund, have economic importance. The accounting has been simplified by neglecting a lot of details. The simplified model is at least internally consistent, and I believe the variables of economic importance are present, and adequately defined.

The computer program which does the calculations is given in Appendix 2. The present appendix will explain the main structure of the program. Although the stochastic price process is in continuous time, the assumption is that the company only performs its activities once a year. The costs are paid for, the petroleum is sold, and financial decisions and tax payments are made, all at the end of each year.

Define TRA_t, TYA_t, TSA_t, TXA_t, TKA_t, and TWA_t as the accrued royalty, income tax, state tax, special petroleum tax, capital tax, and withholding tax, respectively. It is straight
forward to calculate $\text{TRA}_t$. Also, it is straightforward to calculate the book value of assets, $\text{BVA}_t$, each year. The depreciation is always straight-line over six years. If the company has a loss in some year, the excess depreciation is included in the tax-deductible loss to be carried forward. This means that the book value of assets, and the yearly depreciation $\text{DPR}_t = \text{BVA}_{t-1} - \text{BVA}_t + \text{IVS}_t$, only depends on investment costs, not on profits or losses.

Net interest expenditures, $\text{NIE}_t$ (and the non-tax-deductible part, $\text{NDI}_t$), are calculated from a market interest rate and the previous year's net financial assets $\text{CSH}_{t-1} - \text{DEB}_{t-1}$. The profits after royalties, before the other taxes, $\text{PBT}_t$, is given by (1), and if this is negative, then $\text{LOS}_t = -\text{PBT}_t$, otherwise, $\text{LOS}_t = 0$. The uplift allowance, $\text{UPC}_t$, is calculated separately from $\text{BVA}_t$ and $\text{DPR}_t$, since the rules for carry-forward in the case of loss are different. The taxes $\text{TYA}_t$ and $\text{TXA}_t$ can then be calculated directly.

The remaining variables must be determined simultaneously. This has long been acknowledged in the Norwegian accounting literature, see e.g. Vårdal (1984, p. 308), who considers three cases, i.e. three different sets of simultaneous equations, for calculating maximum dividends. His three cases depend on the kind of reserve fund allocations that are made. Considering in addition the non-dividend paying cases, the
maximum debt/equity ratio, and its influence on the capital tax, as many as seventeen cases may occur.

Based on tax rates and the given variables $BVA_t$, $PBT_t$, $LOS_t$, $CCA_t$, $NDI_t$, $TYA_t$, $TXA_t$, and the previous year's $RSF_{t-1}$, and $KST_{t-1}$, the model for each year in the company's life determines the "endogenous" variables $CSH_t$, $DEB_t$, $KST_t$, $RSF_t$, $DIV_t$, $NDD_t$, $TSA_t$, and $TKA_t$, solving a set of eight linear equations. There are eighteen possible equations all together, but only eight apply in each case. Of all possible combinations of these equations, only seventeen may occur in practice, given various constraints on the values of the variables. I will first give the equations with some motivation, and describe the seventeen cases in a table. Then I will discuss the solution method.

Below, equations which are alternative are separated by letters, e.g. (1.2.a) and (1.2.b). The endogenous variables appear on the right-hand side, and the cases in which each equation applies is given in parentheses:

The balance sheet identity,

$$(1.1) \quad BVA_t + LOS_t = DEB_t + KST_t + RSF_t - CSH_t \quad (\text{all cases}).$$

The income statement identity,
\[(1.2.a) \quad \text{PBT}_t - \text{CCA}_t - \text{NDI}_t - \text{TYA}_t - \text{TX}_t - \text{RSF}_{t-1} + \text{LOS}_t \]

\[= \text{TSA}_t + \text{TKA}_t + \text{DIV}_t + \text{NDD}_t + \text{RSF}_t \quad (1 \text{ through } 16). \]

An alternative income statement identity in the case when there is a transfer from the capital stock, \( \text{KTR}_t \), to cover a non-tax-deductible loss which cannot be covered by a reduction in the reserve fund, (in fact, the variable \( \text{KTR}_t \) is a ninth endogenous variable, appearing only in case 17, in which the ninth equation is \( \text{RSF}_t = 0 \)),

\[(1.2.b) \quad \text{PBT}_t - \text{CCA}_t - \text{NDI}_t - \text{TYA}_t - \text{TX}_t - \text{RSF}_{t-1} + \text{LOS}_t \]

\[= \text{TSA}_t + \text{TKA}_t + \text{DIV}_t + \text{NDD}_t + \text{RSF}_t + \text{KTR}_t \quad (17). \]

The constraint on the debt-equity ratio, where \( \text{DEM} \) is the debt/equity maximum,

\[(1.3.a) \quad \text{DEM} \cdot \text{LOS}_t = -\text{DEB}_t + \text{DEM} \cdot (\text{KST}_t + \text{RSF}_t) \]

\[(1,5,6,8,10,13,14,17). \]

The alternative, when the constraint cannot be reached because equity is too large,
(1.3.b) \[ 0 = KST_t \ (2,3,4,7,9,11,12,15,16). \]

The calculation of state and capital taxes,

(1.4) \[ t_s \cdot \text{Max}(PBT_t, 0) = TSA_t + t_s \cdot \text{DIV}_t \quad \text{(all cases)}, \]

and

(1.5) \[ t_k \cdot \text{BVA}_t = TKA_t + t_k \cdot (\text{DEB}_t - \text{CSH}_t) \quad \text{(all cases)}. \]

A binding constraint that \( \text{CSH}_t \) cannot be negative,

(1.6.a) \[ 0 = \text{CSH}_t \quad (1,2,4 \text{ through } 10,13 \text{ through } 17). \]

The alternative, the assumption that there is no debt when there are positive net financial assets, \( \text{CSH}_t - \text{DEB}_t > 0 \), (e.g. because of a positive \( \text{RSF}_t - \text{BVA}_t - \text{LOS}_t \)),

(1.6.b) \[ 0 = \text{DEB}_t \quad (3,11,12). \]

An assumption that non-deductible dividends are zero,

(1.7.a) \[ 0 = \text{NDD}_t \quad (1 \text{ through } 11,13,15,17). \]

Alternatively, tax-deductible dividends reach their maximum,
(1.7.b) \[ PBT_t - NDt = DIV_t \] (12,14,16).

The case of a loss, when dividends are zero by law, and the reserve fund is reduced to cover the non-tax-deductible part of the loss, \( RSF_t \) being determined by the income statement identity (or \( RSF_t = 0 \) in case 17),

(1.8.a) \[ 0 = DIV_t \] (1,2,3,17).

Alternatively, one of the six legal constraints on the reserve fund when there is not a loss, first the two ways in which it may reach its legal minimum exactly,

(1.8.b) \[ 0 = DEB_t - KST_t - RSF_t \] (4),

or

(1.8.c) \[ 0 = FKR*KST_t - RSF_t \] (5),

where FKR is the legal fund-to-capital-stock ratio, currently 20%. Then the two ways in which the fund may be increased without reaching the legal minimum,
(1.8.d) \[ \text{RSF}_{t-1} + \text{OAR} \ast (\text{PBT}_t - \text{TYA}_t - \text{TXA}_t - \text{CCA}_t - \text{NDI}_t) \]

\[ = \text{RSF}_t + \text{OAR} \ast (\text{TSA}_t + \text{TKA}_t) \quad (6,7), \]

where OAR, currently 10\%, is the required ordinary reserve fund allocation rate (out of after-tax profits, when positive), or, when extra-ordinary allocations must be made in addition,

(1.8.e) \[ \text{RSF}_{t-1} + \text{OAR} \ast (\text{PBT}_t - \text{TYA}_t - \text{TXA}_t - \text{CCA}_t - \text{NDI}_t) \]

\[ - \text{EAR} \ast (\text{KST}_{t-1} + \text{RSF}_{t-1}) = \text{RSF}_t + \text{OAR} \ast (\text{TSA}_t + \text{TKA}_t) - \text{DIV}_t \]

\[ (8,9), \]

where EAR, currently 10\%, is the trigger ratio of dividends relative to previous year's KST\(_{t-1}\)+RSF\(_{t-1}\), above which extra-ordinary reserve fund allocations have to be made. Finally, the legal reduction in the reserve fund, when PBT\(_t\) > 0, and the fund exceeds the legal minima, of which the greater should be considered,

(1.8.f) \[ \text{RSF}_{t-1} \ast (1 - \text{RFR}) = \text{RSF}_t - \text{RFR} \ast \text{FKR} \ast \text{KST}_t \quad (10,11,12), \]

where RFR, currently 20\%, is the legal reserve fund reduction rate, or alternatively,
\[(1.8.f) \quad RSF_{t-1}*(1-RFR) = RSF_t-RFR*(KST_t-DEB_t)\]

\[(13,14,15,16).\]

Table 1.1 shows the relation between the seventeen cases that may occur. (ORA and EOA are ordinary and extraordinary reserve fund allocations, respectively.) Cases 1 - 3 have no dividends, either because \(PBT_t<0\), or because a small positive \(PBT_t\), with a possible addition of a reduction in the reserve fund, is not enough to pay both taxes and dividends. Case 17 is an extension of case 1, with \(PBT_t<0\), but with the difference that even though the reserve fund is reduced to zero, this is not enough to cover the non-tax-deductible loss. Since the company is assumed not to write this loss into the balance sheet, part of the capital stock is used to cover the loss.

The cases 5, 10, 13, and 14 can be ruled out when \(DEM>1\), as is the case throughout this study. This leaves 13 possible cases.

For a given set of pre-determined variables, only one of these thirteen cases is the valid solution, but all thirteen systems of eight equations can be solved. In order to pick the right solution, the solutions are checked against inequality
<table>
<thead>
<tr>
<th></th>
<th>$\text{DEB}_t = \text{DEM}^*$</th>
<th>$\text{KST}_t = 0$</th>
<th>$\text{RSF}_t - \text{LOS}_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\text{KST}_t +$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{CSH}_t = 0$</td>
<td>$\text{DD}_t = \text{PBT}_t$</td>
<td>$\text{DD}_t = \text{PBT}_t$</td>
<td>$\text{DD}_t = \text{PBT}_t$</td>
</tr>
<tr>
<td></td>
<td>$= \text{NDI}_t$</td>
<td>$= \text{NDI}_t$</td>
<td>$= \text{NDI}_t$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{DIV}_t = 0$</td>
<td>1, 17</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$\text{RSF}_t = \text{DEB}_t - \text{KST}_t$</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{RSF}_t = \text{FKR}^* \text{KST}_t$</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{RSF}<em>t = \text{RSF}</em>{t-1} + \text{ORA}_t$</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{RSF}<em>t = \text{RSF}</em>{t-1} + \text{ORA}_t + \text{EOA}_t$</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{RSF}<em>t = \text{RSF}</em>{t-1}$</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>$= \text{RSF}_t - \text{FKR}^* \text{KST}_t$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{RSF}<em>t = \text{RSF}</em>{t-1}$</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>$= \text{RSF}_t - \text{DEB}_t + \text{KST}_t$</td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.1: The seventeen different cases of combinations of binding constraints on the year-end variables.

Constraints. I will not write down all of them here. They can be sorted in three groups. First, all variables have to be positive. Secondly, there are constraints on maximum debt/equity ratio and maximum tax-deductible dividends, the latter being $\text{PBT}_t - \text{NDI}_t$. 
Thirdly, there are the rules that determine whether and how the reserve fund should be decreased or increased. For instance, the cases 4 and 5, in which the reserve fund reaches its legal minimum exactly, can not occur if $RSF_t < RSF_{t-1}$, and they are assumed only to occur if the increase in the fund has been less than or equal to what the formulae for ordinary and extraordinary allocations prescribe. (In the year in which the legal minimum RSF is met, the company does not have to make full allocations.) Another example: The cases 13 through 16, in which the $RSF_t$ is reduced by 20% of its excess over $DEB_t - KST_t$, can only be valid solutions if $DEB_t - KST_t$ exceeds 20% of $KST_t$, and if at the same time $RSF_t$ exceeds $DEB_t - KST_t$. Such complicated rules imply that the problem cannot be formulated as a standard linear programming problem. The computer program solves one case, i.e. a set of equations, then checks whether the solution is valid. If not, it continues with other cases. I believe this is computationally as efficient as determining in which areas of the space of pre-determined variables the different cases are valid.

It will not be proved here that the list of cases is exhaustive, nor that there is only one solution for each set of pre-determined variables. That the list is in fact exhaustive is indicated very strongly from the simulations. For more than one million different sets of pre-determined
variables, there has always been a solution among the thirteen. There are, however, sets of pre-determined variables for which there is more than one solution, if one allows for the possibility that the company chooses not to pay out dividends even if it can. The first version of the computer model checked all thirteen cases for each set of pre-determined variables, and on some occasions, both a case-3 solution and a solution with dividend payout would satisfy all constraints. A faster version of the program only checks cases until one solution is found, and this version checks dividend-paying cases first, on the assumption that companies will prefer to pay dividends if they can.

Table 1.2 shows the occurrence of the various cases in 5000 runs of the model for the before-tax optimal one-field development, with a life-time of 23 years. The first and the last year are not counted, as the calculations are different for those years. The numbers in the table thus sum to 105000.

In order to speed up computations, the order of the cases in the computer program is such that within the dividend-paying cases, the most frequent cases are checked first, and similarly within the non-dividend-paying cases.

The table shows clearly the importance of the reserve fund requirement. The slow legal build-down of the fund when it
Table 1.2: The frequency of occurrence of the seventeen cases in 5000 runs of the simulation model, with the before-tax optimal one-field development.

<table>
<thead>
<tr>
<th>Case</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq.</td>
<td>22889</td>
<td>390</td>
<td>1316</td>
<td>1574</td>
<td>0</td>
<td>889</td>
<td>29</td>
<td>4850</td>
<td>1769</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq.</td>
<td>0</td>
<td>15049</td>
<td>1368</td>
<td>0</td>
<td>0</td>
<td>4622</td>
<td>263</td>
<td>49992</td>
</tr>
</tbody>
</table>

exceeds the minimum leads to frequent occurrence of the cases 10 through 16 in which the fund is built down. A too large fund is the only thing that can force the companies to hold positive net financial assets. This happens in case 11, which occurs very frequently.
Appendix 2:  
Simulation Computer Program.
The computer program was written in the language GAUSS, 
version 1.3B. The program is reproduced below. Before 
running the program, the total number of runs must be given in 
the variable TOTAL, and the random number generator may be 
seeded. Explanation of variable names are given at the end.

(Beginning of program.)
FORMAT /M1 /RD 9,0;CLS;
@ Save start time. @
TEMPOR=DATE;STADAT=TEMPOR[1:3,,]’;
TEMPOR=TIME;STATIM=TEMPOR[1:3,,]’;
@ Due to precision problems, an additive fudge factor of
0.0000000001 is allowed in inequality constraints. @
FDG=0.0000000001;
@ Load scalar data from the file SCALAR, and assign them. @
LOAD SCALAR[21,1]=SCALAR.DTA;RFI=SCALAR[1,1];RFR=SCALAR[2,1];
DEM=SCALAR[3,1];OAR=SCALAR[4,1];EAR=SCALAR[5,1];
Tyr=SCALAR[6,1];TSR=SCALAR[7,1];TXR=SCALAR[8,1];
TRW=SCALAR[9,1];TKR=SCALAR[10,1];NFI=SCALAR[11,1];
THO=SCALAR[12,1];TRB=SCALAR[13,1];TRG=SCALAR[14,1];
MNP=SCALAR[15,1];Fy=SCALAR[16,1];IFL=SCALAR[17,1];
STL=SCALAR[18,1];CYO=SCALAR[19,1];PRI1=SCALAR[20,1];
FKR=SCALAR[21,1];
@ Load matrix data. @
LOAD SDE[MNP,NFI]=SDE.DTA;LOAD FIE[NFI,6]=FIE.DTA;
LOAD CMX[128,8]=CMX.DTA;
@ Load quantities, operating costs, investments. @
NFI2=NFI*2;NFI*MNP;
LOAD QUA[THO,NFI2]=QUA.DTA;LOAD OPC[THO,NFI]=OPC.DTA;
LOAD IVS[THO,NFI*MNP]=IVS.DTA;LOAD CCB[THO,NFI]=CCB.DTA;
@ Find completion dates of investment for each field. @
I=1;
DO UNTIL I>NFI;
  CDIF=THO+1
    MININDC(REV(SUMC(IVS[,(I-1)*MNP+1:I*MNP]').==0));
    IF I==1;CDI=CDIF;ELSE;CDI=CDI;CDIF;ENDIF;
I=I+1;
ENDO;
@ Initialize price independent matrices. @
BV=ZEROS(1,6*NFI);TRO=ZEROS(1,NFI);CS$ESAV=ZEROS(18,1);
CLEAR BVA,DRP,CFA,SUMT,SUMAT,SUMBTSQ,SUMATSQ;
@ Initially, all abandons are set to the last period. @
ABD=THO*ONES(NFI,1);
@ Loop runs once for each time period from 2 up to THO. @
T=2;
DO UNTIL T>THO;
  TLI=T-1;CCA=CCA:0; @ (Initially.) @
  @ Loop runs once for each field. @
  I=1;
DO UNTIL I>NFI;
  @ Calculate royalty rates for oil. @
  QUAOIL=QUA[T,2*I-1];
  IF FIE[I,1]==1;TROC=0;
  ELSE;
    IF QUAOIL>FIE[I,6];TROC=0.16;
    ELSEIF QUAOIL>FIE[I,5];TROC=0.14;
    ELSEIF (QUAOIL>FIE[I,4] OR TRO[TL1,I]>0.11);
    TROC=0.12;
    ELSEIF QUAOIL>FIE[I,3];TROC=0.1;
    ELSE;TROC=0.08;
  ENDIF;
ENDIF;
IF I=1;TROALL=TROC;ELSE;TROALL=TROALL~TROC;ENDIF;
  @ Calculate possible abandonments of fields. @
IF T>CDI[1,1] AND T<ABD[I,1] AND
  QUA[MINC(T+1:THO),2*I-1:2*I]<0;
  ABD[I,1]=T;OPC[T+1:THO,I]=ZEROS(THO-T,1);
  QUA[T+1:THO,2*I-1:2*I]=ZEROS(THO-T,2);
  CCA[T,1]=CCA[T,1]+CCB[T,I]*(1-TYR-TSR-TRXR);
  CCB[T+1:THO,I]=ZEROS(THO-T,1);
ELSEIF T==THO;
  CCA[T,1]=CCA[T,1]+CCB[T,I]*(1-TYR-TSR-TRXR);
ENDIF;
  @ Calculate book value of assets per field. @
IF ABD[I,1]<T;BVC=ZEROS(1,6);
ELSE;
  FC6=6*(I-1);FCM=MNP*(I-1)+1;
  BVC=BV[TL1,FC6+1]
    -SUMC(SUMC(IVS[1:TL1,FCM:MNP*I]).*(SDE[.,1].==T))
    +SUMC(SUMC(IVS[T,FCM:MNP*I]).*(SDE[.,1].>T))
    *(5/6)^*(SUMC(IVS[T,FCM:MNP*I]).*(SDE[.,I].<=T))
    +SUMC(SUMC(IVS[1:TL1,FCM:MNP*I]).*(SDE[.,I].==T))
    *(0.8*BV[TL1,FC6+2]-0.75*BV[TL1,FC6+3]
    -(2/3)*BV[TL1,FC6+4]-0.5*BV[TL1,FC6+5];
ENDIF;
IF I=1;BV=BVC;ELSE;BV=BV+BVC;ENDIF;
I=I+1;
ENDO;
BV=BV;BVV;TRO=TRO;TROALL;
  @ Calculate book value of assets, and depreciation. @
BVA=BVA+SUMC(BVV');
DPR=DPR+BVA[TL1,1]-BVA[T,1]+SUMC(IVS[T,..]');
T=T+1;
ENDO;
@ Start simulation. Variable TOTAL is initialized
  before program is started. @
NRN=1;
DO UNTIL NRN>TOTAL;
  @ Initialize price-dependent matrices.
Gas price = 0.8 * oil price. @
PRI=PRI1*(1.0.8);UP=ZEROS(1,15);CLEAR KST,CSH,CSE,DEB,RSF;
CLEAR LOS,NIE,TRA,TYA,TSA,TXA,TKA,TWA,PBT,DIV,NDD,ND1,RV;
CLEAR UPU,NIU,KTR;
@Loop runs once for each time period from 2 up to THO. @
T=2;
DO UNTIL T>THO;
   TL=T-1;
   @ Random oil price and gas price. @
   PRI=PRI1*PRI[TL,]*EXP(RNDN(1,1)*STL+IFL+RFI-CYO-STL^2/2);
   @ Loop calculates royalty for each field. @
   I=1;
   DO UNTIL I>NFI;
      IF FIE[I,1]=1;
         TRAC=TRB*(PRI[T,1]*QUA[T,2*I-1]+PRI[T,2]*QUA[T,2*I]);
      ELSE;TRAC=TRG*(PRI[T,2]*QUA[T,2*I])
            +TRO[T,1]*(PRI[T,1]*QUA[T,2*I-1]);
      ENDF;
      IF I=1;TRA=TRA;TRAC;ELSE;TRA[T,1]=TRA[T,1]+TRAC;ENDIF;
      I=I+1;
   ENDO;
   @ Uplift for future years, maximum uplift for current yr. @
   NIU=NIU1.2*SUMC(SUMC(MAT(BV,T,SEQA(2,6,NFI)))
   IF MAXC(ABD)<T;UPC=SUMC(UP[TL,]+NIU[T,1];
   UP=UP;ZEROS(1,15);
   ELSE;UPC=UP[TL,1];UP=UP:(NIU[T,1]15+(UP[TL,1215]0));
   ENDF;
   @ Calculate net interest expenditures. @
   IF SUMC(SUMC(MAT[1:T,]))>0;
      NIE=NIE:(DEB[TL,1]-CSH[TL,1])*(EXP(RFI+IFL)-1);
      ND1=ND1;
   ELSE;ND1=ND1:(DEB[TL,1]-CSH[TL,1])*(EXP(RFI+IFL)-1);
      NIE=NIE0;
   ENDF;
   @ Calculate profits after royalties, before taxes. @
   RVN=RVN1PRI[T,1]*SUMC(RESHAPEDQUA[T,1],NFI,2));
   PBT=PBT1RVN[T,1]-TRA[T,1]-SUMC(OPC[T,1])
      -NIE[T,1]-DPR[T,1]-LOS[TL,1];
   @ Calculate income tax. Jump to final year's calculations
   if all fields are abandoned and final year's calculations
   are desired. If not, calculate special petroleum tax,
   and adjust next year's uplift. Calculate the
tax-deductible loss to be carried forward. @
   TYA=TYA+TYR*MAXC(PBT[T,1];0);
   IF FY=1 AND T=MAXC(ABD);GOTO FINALYR;ENDIF;
   TX=TXA;TXR*MAXC(PBT[T,1]-UPC;0);UP=UPU
      ;MAXC(0;MINC(PBT[T,1];UPC));LOS=LOS;MAXC(-PBT[T,1];0);
   UP[T,1]=UP[T,1]+MAXC(0;MINC(UPC-PBT[T,1];UPC));
   @ Solve the systems of simultaneous linear equations.
   First try the 15 cases in which there is no use of the
   capital stock to cover losses, either because there is no
   non-tax-deductible loss, or because RSF covers it. @
BMTX = @1@ BVA[T,1]+LOS[T,1]; @2@ PBT[T,1]-CCA[T,1]-NDI[T,1]
-TYA[T,1]-TXA[T,1]+RSF[TL1,1]+LOS[T,1]; @3@ DEM*LOS[T,1]
@4@ 0; @5@ TSR*MAXC(0;PBT[T,1]); @6@ TKR*BVA[T,1]; @7@ 0;
@8@ 0; @9@ 0; @10@ PBT[T,1]-NDI[T,1]; @11@ 0; @12@ 0;
@13@ 0; @14@ RSF[TL1,1]+OAR*(PBT[T,1]-TYA[T,1]-TXA[T,1]
-CCA[T,1]-NDI[T,1]); @15@ RSF[TL1,1]+OAR*(PBT[T,1]
-TYA[T,1]-TXA[T,1]-CCA[T,1]-NDI[T,1]-EAR*(KST[TL1,1]+
RSF[TL1,1])); @16@ RSF[TL1,1]*(1-RFR); @17@ RSF[TL1,1]
*(1-RFR);
SOL=ZEROS(9,1);
IF LOS[T,1]>FDG;GOTO LOSS;ENDIF;
XXX=CMX[81:88,.]*SUBMAT(BMTX,1:2:4:5:6:8:9:16,0);
IF XXX[1,1]>=(-FDG);
IF XXX[7,1]<=(-FDG);
IF XXX[4,1]<>FDG+PBT[T,1]-NDI[T,1];
IF XXX[4,1]<>FDG;
IF XXX[4,1]<RSF[TL1,1]-FDG;
SOL=SOL~(XXX:11);GOTO ASSIGN;ENDIF;ENDIF;ENDIF;ENDIF;ENDIF;
XXX=CMX[57:64,.]*SUBMAT(BMTX,1:2:3:5:6:7:9:15,0);
IF XXX[3,1]<>FDG;
IF XXX[4,1]<FDG+MAXC(FKR*XXX[3,1];XXX[2,1]-XXX[3,1]);
IF XXX[7,1]<(-FDG)+EAR*(KST[TL1,1]+RSF[TL1,1]);
IF XXX[7,1]<FDG+PBT[T,1]-NDI[T,1];
SOL=SOL~(XXX:8);GOTO ASSIGN;ENDIF;ENDIF;ENDIF;ENDIF;ENDIF;
XXX=CMX[113:120,.]*SUBMAT(BMTX,1:2:4:5:6:7:9:17,0);
IF XXX[2,1]<>FDG;
IF XXX[2,1]<FDG+DEM*(XXX[3,1]+XXX[4,1]-LOS[T,1]);
IF XXX[4,1]>XXX[2,1];
IF XXX[7,1]<(-FDG);
IF XXX[7,1]<FDG+PBT[T,1]-NDI[T,1];
IF XXX[4,1]<RSF[TL1,1]-FDG;
SOL=SOL~(XXX:15);GOTO ASSIGN;ENDIF;ENDIF;ENDIF;ENDIF;ENDIF;
XXX=CMX[65:72,.]*SUBMAT(BMTX,1:2:4:5:6:7:9:15,0);
IF XXX[4,1]<FDG+MAXC(FKR*XXX[3,1];XXX[2,1]-XXX[3,1]);
IF XXX[7,1]<(-FDG)+EAR*(KST[TL1,1]+RSF[TL1,1]);
IF XXX[2,1]<>FDG;
IF XXX[2,1]<FDG+DEM*(XXX[3,1]+XXX[4,1]-LOS[T,1]);
IF XXX[7,1]<FDG+PBT[T,1]-NDI[T,1];
SOL=SOL~(XXX:9);GOTO ASSIGN;ENDIF;ENDIF;ENDIF;ENDIF;ENDIF;
XXX=CMX[89:96,.]*SUBMAT(BMTX,1:2:4:5:6:8:10:16,0);
IF XXX[1,1]<>FDG;
IF XXX[8,1]<>(-FDG);
IF XXX[4,1]<>(-FDG);
IF XXX[4,1]<RSF[TL1,1]-FDG;
SOL=SOL~(XXX:12);GOTO ASSIGN;ENDIF;ENDIF;ENDIF;ENDIF;
XXX=CMX[25:32,.]*SUBMAT(BMTX,1:2:4:5:6:7:9:12,0);
IF XXX[2,1]<>FDG;
IF XXX[2,1]<FDG+DEM*(XXX[3,1]+XXX[4,1]-LOS[T,1]);
IF XXX[4,1]<RSF[TL1,1]<FDG+OAR*(PBT[T,1]-TYA[T,1]
-TXA[T,1]-XXX[5,1]-XXX[6,1]-CCA[T,1]-NDI[T,1])
+\text{MAXC}(0:XXX[7,1]) - \text{EAR}((\text{KST}[\text{TL1},1] + \text{RSF}[\text{TL1},1]));
\text{IF} \ XXX[7,1] > (\text{FDG});
\text{IF} \ XXX[4,1] = (\text{FDG}) + \text{RSF}[\text{TL1},1];
\text{SOL} = \text{SOL}^\sim(\text{XXX}:4); \text{GOTO} \ \text{ASSIGN}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF};
\text{XXXX} = \text{CMX}[41:48,.] \ast \text{SUBMAT}(\text{BMTX},1:2:3:5:6:7:9:14,0);
\text{IF} \ XXX[7,1] < (\text{FDG} + \text{EAR}((\text{KST}[\text{TL1},1] + \text{RSF}[\text{TL1},1]));
\text{IF} \ XXX[3,1] < (\text{FDG});
\text{IF} \ XXX[4,1] < (\text{FDG}) + \text{MAXC}((\text{FKR} * \text{XXX}[3,1] ; \text{XXX}[2,1] - \text{XXX}[3,1]));
\text{IF} \ XXX[7,1] > (\text{FDG});
\text{IF} \ XXX[7,1] < (\text{FDG} + \text{PBT}[T,1] - \text{NDI}[T,1]);
\text{SOL} = \text{SOL}^\sim(\text{XXX}:6); \text{GOTO} \ \text{ASSIGN}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF};
\text{XXXX} = \text{CMX}[121:128,.] \ast \text{SUBMAT}(\text{BMTX},1:2:4:5:6:7:10:17,0);
\text{IF} \ XXX[2,1] > (\text{FDG});
\text{IF} \ XXX[4,1] < (\text{FDG} + \text{DEM}((\text{XXX}[3,1] + \text{XXX}[4,1] - \text{LOS}[T,1]));
\text{IF} \ XXX[4,1] > \text{XXX}[2,1];
\text{IF} \ XXX[8,1] > (\text{FDG});
\text{IF} \ XXX[4,1] < \text{RSF}[\text{TL1},1] - \text{FDG};
\text{SOL} = \text{SOL}^\sim(\text{XXX}:16); \text{GOTO} \ \text{ASSIGN}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF};
\text{LOSS:}
\text{XXXX} = \text{CMX}[1:8,.] \ast \text{SUBMAT}(\text{BMTX},1:2:3:5:6:7:9:11,0);
\text{IF} \ XXX[3,1] > (\text{FDG});
\text{IF} \ XXX[4,1] = (\text{FDG} + \text{RSF}[\text{TL1},1];
\text{IF} \ XXX[4,1] > (\text{FDG});
\text{SOL} = \text{SOL}^\sim(\text{XXX}:1); \text{GOTO} \ \text{ASSIGN}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF};
\text{XXXX} = \text{CMX}[17:24,.] \ast \text{SUBMAT}(\text{BMTX},1:2:4:5:6:8:9:11,0);
\text{IF} \ XXX[1,1] > (\text{FDG});
\text{IF} \ XXX[4,1] < (\text{FDG} + \text{RSF}[\text{TL1},1];
\text{IF} \ XXX[4,1] > (\text{FDG});
\text{SOL} = \text{SOL}^\sim(\text{XXX}:3); \text{GOTO} \ \text{ASSIGN}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF};
\text{XXXX} = \text{CMX}[9:16,.] \ast \text{SUBMAT}(\text{BMTX},1:2:4:5:6:7:9:11,0);
\text{IF} \ XXX[2,1] > (\text{FDG});
\text{IF} \ XXX[2,1] < (\text{FDG} + \text{DEM}((\text{XXX}[3,1] + \text{XXX}[4,1] - \text{LOS}[T,1]));
\text{IF} \ XXX[4,1] < (\text{FDG} + \text{RSF}[\text{TL1},1];
\text{IF} \ XXX[4,1] > (\text{FDG});
\text{SOL} = \text{SOL}^\sim(\text{XXX}:2); \text{GOTO} \ \text{ASSIGN}; \text{ENDIF}; \text{ENDIF}; \text{ENDIF};
\text{IF} \ \text{DEM} \ 1 \ \text{AND} \ \text{COLS(SOL)} = 1; \text{GOTO} \ \text{CASE}17; \text{ENDIF};
\text{XXXX} = \text{CMX}[33:40,.] \ast \text{SUBMAT}(\text{BMTX},1:2:3:5:6:7:9:13,0);
\text{IF} \ XXX[3,1] > (\text{FDG});
\text{IF} \ XXX[2,1] < \text{XXX}[3,1] < (\text{FDG} + \text{FKR} * \text{XXX}[3,1];
\text{IF} \ XXX[4,1] = \text{RSF}[\text{TL1},1] = (\text{FDG} + \text{OAR}((\text{PBT}[T,1] - \text{TYA}[T,1])}
-TXA[T,1]-XXX[5,1]-XXX[6,1]-CCA[T,1]-NDI[T,1])
+MAXC(0;XXX[7,1]-EAR*(KST[TL1,1]+RSF[TL1,1]));
IF XXX[7,1]>(-FDG);
IF XXX[7,1]<=FDG+PBT[T,1]-NDI[T,1];
IF XXX[4,1]>(-FDG)+RSF[TL1,1];
SOL=SOL*(XXX:5);GOTO ASSIGN;ENDDIF;ENDDIF;ENDDIF;ENDDIF;ENDDIF;ENDDIF;
ENDIF;
XXX=CMX[73:80,..]*SUBMAT(BMTX,1:2:3:5:6:7:9:16,0);
IF XXX[3,1]>(-FDG);
IF XXX[2,1]-XXX[3,1]<=FDG+FKR*XXX[3,1];
IF XXX[4,1]>FKR*XXX[3,1];
IF XXX[7,1]>(-FDG);
IF XXX[7,1]<=FDG+PBT[T,1]-NDI[T,1];
IF XXX[4,1]<=RSF[TL1,1]-FDG;
SOL=SOL*(XXX:10);GOTO ASSIGN;ENDDIF;ENDDIF;ENDDIF;ENDDIF;ENDDIF;ENDDIF;
ENDIF;
XXX=CMX[97:104,..]*SUBMAT(BMTX,1:2:3:5:6:7:9:17,0);
IF XXX[2,1]-XXX[3,1]>FKR*XXX[3,1];
IF XXX[4,1]>XXX[2,1]-XXX[3,1];
IF XXX[7,1]>(-FDG);
IF XXX[7,1]<=FDG+PBT[T,1]-NDI[T,1];
IF XXX[4,1]<=RSF[TL1,1]-FDG,
SOL=SOL*(XXX:13);GOTO ASSIGN;ENDDIF;ENDDIF;ENDDIF;ENDDIF;ENDDIF;ENDDIF;
XXX=CMX[105:112,..]*SUBMAT(BMTX,1:2:3:5:6:7:10:17,0);
IF XXX[2,1]-XXX[3,1]>FKR*XXX[3,1];
IF XXX[4,1]>XXX[2,1]-XXX[3,1];
IF XXX[8,1]>(-FDG);
IF XXX[4,1]<=RSF[TL1,1]-FDG;
SOL=SOL*(XXX:14);GOTO ASSIGN;ENDDIF;ENDDIF;ENDDIF;ENDDIF;ENDDIF;
GOTO CASE17;

ASSIGN:
@ Unique solution found among first 16 cases. @
SOL=SOL[1,2];CASH=CASH:SOL[1,1];DEB=DEB:SOL[2,1];
KST=KST:SOL[3,1];RSF=RSF:SOL[4,1];TSA=TSA:SOL[5,1];
TKA=TKA:SOL[6,1];DIV=DIV:SOL[7,1];NDD=NDD:SOL[8,1];
CSE=CSE:SOL[9,1];TWA=TWA;TWR*(DIV[T,1]+NDD[T,1]);KTR=KTR:0;
GOTO YEAREND;

CASE17:
@ Current year's non-tax-deductible loss exceeds RSF, and
some of the capital stock must be used to cover the loss,
perhaps out of new share issues. @
TSATMP=TSR*MAXC(0;PBT[T,1]);TKATMP=TKR*BVA[T,1]/(1+DEM);
KTRTMP=-(PBT[T,1]-CCA[T,1]-NDI[T,1]-TYA[T,1]-TXA[T,1]
-TSATMP- TKATMP+RSF[TL1,1]+LOS[T,1]);
IF KTRTMP>FDG;
    TSA=TSA;TSATMP;TKA=TKA;TKATMP;KTR=KTR;KTRTMP;CASH=CASH:0;
    NDD=NDD:0;DIV=DIV:0;RSF=RSF:0;TW=TWA:0;
    DEB=DEB:BVA[T,1]*DEM/(1+DEM);KST=KST:LOS[T,1]
    +BVA[T,1]/(1+DEM);CSE=CSE:17;GOTO YEAREND;
ENDIF;
"No cases passed the tests.";END;
YEAREND:
  T=T+1;
ENDO;
"No final-year calculations have been made.";END;
FINALLYR:
CSH=CSH:0;DEB=DEB:0;KST=KST:0;TKA=TKA:0;RSF=RSF:0;LOS=LOS:0;
CSE=CSE:0;THO=T;
@ Carry back uplift if necessary. @
IF PBT[T,1]>UPC;TXA=TXA:TXR*(PBT[T,1]-UPC);UPU=UPU:UPC;
ELSE;
  UPB=MINC(UPC:UPC-PBT[T,1]:SUMC((NIU[1:T-2,1]:0)
        -UPU[1:T-1,1]));
  UPR=MAXC(PBT[T,1]:0);TB=T-1;
  DO UNTIL TB<3;
    UPR=MINC(UPB[1,1]:TXA[TB,1]/TXR):UPR;
    UPB=MINC(UPB[1,1]-UPR[1,1]:SUMC((NIU[1:TB-2,1]:0)
        -UPU[1:TB-1,1]));UPB;
  TB=TB-1;
ENDO;
  UPR=0;MINC(UPB[1,1]:TXA[2,1]/TXR):UPR;UPB=0:0:UPB;
  UPU=UPU:SUMC(UPR);TXA=TXA:((-TXR*(UPU[T,1]-UPR[T,1])));
ENDIF;
@ Assign as FYS the final-year surplus before TSA. @
FYS=PBT[T,1]-TYA[T,1]-TXA[T,1]-CCA[T,1]-NDI[T,1]+CSH[TL1,1]
    -DEB[TL1,1]+BVA[TL1,1]+LOS[TL1,1]-KST[TL1,1];
IF FYS<0;
  KTR=KTR:((-FYS);TSA=TSA:TSR*MAXC(PBT[T,1]:0);DIV=DIV:0;
    NDD=NDD:0;TWA=TWA:0;
ELSE;
  KTR=KTR:0;
  IF PBT[T,1]>0;
    IF PBT[T,1]>((FYS-TSR*PBT[T,1])/(1-TSR);
      NDD=NDD:0;DIV=DIV:((FYS-TSR*PBT[T,1])/(1-TSR);
    ELSE;
      NDD=NDD:FYS-PBT[T,1];DIV=DIV:PBT[T,1];
    ENDIF;
    TSA=TSA:TSR*(PBT[T,1]-DIV[T,1]);
ELSE;
    DIV=DIV:0;NDD=NDD:FYS;TSA=TSA:0;
  ENDIF;
  TWA=TWA:TWR*(DIV[T,1]+NDD[T,1]);
ENDIF;
@ Save frequency of cases. @
CN=1;
DO UNTIL CN>THO;
  CSESAV[CSE[CN,1]+1,1]=CSESAV[CSE[CN,1]+1,1]+1;CN=CN+1;
ENDO;
@ Calculate present values. @
DELKST=(KST[1,1];(KST[2:THO-1,1]-KST[1:THO-2,1]):0)+KTR;
CFM=SUMC(((-IVS~OPC~CCA/(1-TYR-TSR-TXR))~RVN)’)
   ~(-DELKST+DIV+NDD-TWA);
NPVMAT=CFM.*SEQM(1,REXPF=RFI-IFL),THO;NPVSUM=SUMC(NPVMAT);
SUMBT=SUMBT+NPVSUM[1,1];SUMBTSQ=SUMBTSQ+NPVSUM[1,1]^2;
SUMAT=SUMAT+NPVSUM[2,1];SUMATSQ=SUMATSQ+NPVSUM[2,1]^2;
SUMTASQ=SUMTASQ+(NPVSUM[1,1]-NPVSUM[2,1])^2;
@ Message to screen. @
IF NRNS>1;
    LOCATE 1,1;FORMAT 9,0,"Completed run number ";NRNS;
    "SUMBT, SUMBTSQ, SUMAT, SUMATSQ, SUMTA, SUMTASQ.";
    SUMBT SUMBTSQ;SUMAT SUMATSQ;SUMTA SUMTASQ;
ENDIF;
@ Save results after every 100'th run. @
IF FLOOR(NRNS/100)=NRNS/100;
    TEMPOR=DATE;TMPDAT=TEMPOR[1,3,.]';
    TEMPOR=TIME;TMPTIM=TEMPOR[1,3,.]';
    FORMAT /M1 /RD 5,0;
    OUTPUT FILE=B:SUM.SAV RESET;
    NRNS;
    STATDAT^STATIM^TMPDAT^TMPTIM;
    CSESAV';
    FORMAT 30,14;
    SUMAT^SUMATSQ;
    SUMBT^SUMBTSQ;
    SUMBT-SUMAT^SUMTASQ;
    OUTPUT OFF;
ENDIF;
NRNS=NRNS+1;
ENDO;
@ Save stop time. @
TEMPOR=DATE;STODAT=TEMPOR[1,3,.]';
TEMPOR=TIME;STOTIM=TEMPOR[1,3,.]';
(End of program.)

Explanatory Notes,
The tax rates t_1, t_2, etc., are named TYR, TSR, etc. THO is
lifetime of company in years, NFI is number of fields, MNP is
maximum number of platforms per field. The input data are
given in nine files. QUA, OPC, IVS, CCB all have THO rows,
QUA has 2*NFI columns, to include gas production, IVS has
columns for the maximum number of platforms. Platforms can
only be written down from their SDE, start date of extraction,
given in a separate file. FIE contains data on company's
share, product type, and resulting royalty schedule for each
field. CMX contains the 16 8x8 coefficient matrices of the
simultaneous equation systems, inverted, for cases 1-16.

ABD = abandonment date,
BMTX = constant terms in simultaneous equations,
BV = book values dated according to depreciation start,
BVC = new dated book values for field I,
BVV = vector of new dated book values for all fields,
CCB = potential clean-up costs before refund,
CDI, CDIF = completion dates for investment,
CFM = cash flow matrix,
CSE, CSESAV = summary statistics of occurrence of 17 cases,
CYO = convenience yield of oil,
FIY = toggle for final-year calculations,
IFL = yearly inflation rate,
NIE = net interest expenditures, tax deductible,
NIU = new investments for uplift, i.e. investment with depreciation start one year earlier,
PRI = oil and gas prices,
PRI1 = start value for oil price,
RFI = risk free interest rate,
RFR = reserve fund reduction rate,
RVN = revenue,
SOL = accepted solution to equation system,
STL = standard deviation of \( \ln \left( \frac{P_t}{P_{t-1}} \right) \),
T, TB = time indices,
TRAC = current accrued royalty for field I,
TRG, TRB, TRO = royalty rates for gas, and oil before and after 1972,
TROALL = royalty rates, all fields, oil after 1972,
TROC = royalty rates oil after 1972 for field I,
UP = uplift for fifteen future years,
UPB = uplift to carry back,
UPC = maximum uplift for current year,
UPR = uplift refunded,
UPU = effectively deducted uplift,
XXX = potential solution to equation system.
CHAPTER 3.

INVESTING IN NON-MARKETABLE ASSETS.
1. Introduction.

Models of capital markets under uncertainty describe the valuation of risky assets in equilibrium. This literature does not say much about the valuation of assets that are not traded in capital markets.¹ Most of the literature explicitly or implicitly assumes that all assets are traded. But there are some exceptions, e.g. Mayers (1972) and (1973), Brito (1977), and Grossman and Shiller (1982). These authors consider the valuation of marketable (traded) assets in economies where not all assets are marketable. The most important non-marketable asset type is assumed to be human capital.

The present paper will discuss the valuation of non-marketable assets: When will an investor with a specified objective function want to possess more of a non-marketable asset? This is a well-defined problem, and it is possible to characterize the investment criterion, based on the same portfolio theory that underlies the market equilibrium models. It is obvious that such investments are undertaken in the real world, since without them, non-marketable assets would hardly exist. Investment in human capital through education is one example.

The investment criterion here is a generalization of previous models, valid irrespective of the ability of the investor to compose an optimal portfolio of risky assets.
The Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) is based on the assumption that investors only care about the mean and the variance of the return on their portfolios. A main result is that the required expected return on an asset is explained by the return's covariance with the return on the market portfolio, and not e.g. by the variance of the return. The model below will show that even under much more general assumptions, the investment criterion depends on a covariance measure. This result does not depend on the mean-variance objective function.

Section two derives an investment criterion for a non-marketable asset in a two-period model. Section three discusses the extention to a multi-period model. Section four looks at the relevance of financing and reinvestment patterns for the investment decision, a much-discussed problem in public finance. Invoking the mean-variance assumption, section five draws the connection to the work by Mayers, and criticizes an application of this in the accounting literature. Section six discusses the applicability of the results, and section seven draws conclusions.
2. The Two-Period Model.

In this section it will be shown that even if an investor is unable to compose an optimal portfolio, the relevant risk measure for an investment is the covariance between its return and marginal utility of consumption. In order to derive the investment criterion, I will set up a simple portfolio model in which the holdings of some assets are effectively constrained from below. The shadow prices of the constraints will be used to derive the criterion for investing in more of these assets. When a constraint is binding, the investor has "too much" of that asset at the going market price. When will you want to buy more if you have too much? If you can get additional amounts of the asset cheap enough.

The case in which some asset holdings are constrained from above is left as an exercise for the reader. The investment criterion will be of exactly the same form. The investor will in that case be willing to pay more than the market price for additional amounts of the asset.

Constrained asset holdings is only one possible formulation of the problem. It suggests that the asset may be traded in the market, but that this particular investor is not able to trade it freely. An alternative formulation is that the individual gets some unique random income in the future, and has some limited opportunity to choose the quantity of this
(probability distribution of) income. This formulation, which gives the same investment criterion, will be used in section five.

The investor has a time-additive\textsuperscript{3} expected utility function over the two\textsuperscript{4} periods, and solves

\begin{equation}
\max_{X_0, \ldots, X_n} u(W_0 - \sum_{j=0}^{n} X_j P_{j0}) + \theta \mathbb{E}(u(\sum_{j=0}^{n} X_j P_{j0} R_j)),
\end{equation}

s.t. $X_j \leq \bar{X}_j$, $j = 0, \ldots, n$,

where $u$ is a concave, increasing, per-period utility function, $W_0$ is initial wealth, $X_j$ is the quantity held of asset $j$, $j=0, \ldots, n$, $P_{j0}$ is the price of asset $j$ at time zero, $P_{j1}$ is its random price\textsuperscript{5} at time one, and $R_j = P_{j1}/P_{j0}$ is its random return (one plus the rate of return) from time zero to time one. The investor takes the probability distributions of prices and returns as exogenously given. $\mathbb{E}$ denotes expectation. $\theta$ is the utility discount factor. The argument of the first $u$ function will be called $C_0$, while the second is $C_1$.\textsuperscript{6} It is assumed that $W_0$ exceeds the sum of all $\bar{X}_j P_{j0}$, such that the investor is able to obtain a positive $C_0$, and that any positive consumption is preferred to negative consumption.

The Lagrangian of the problem is
\[ L = u(W_0 - \sum_{j=0}^{n} X_j P_{j0}) + P_{j0} \sum_{j=0}^{n} \lambda_j (X_j - \bar{X}_j), \]

and the first-order conditions for a maximum are, for \( j = 0, \ldots, n, \)

\[ \frac{\partial L}{\partial X_j} = u'(C_0)(-P_{j0}) + \theta E(u'(C_1)P_{j1}) + \lambda_j = 0, \]

or \( X_j = 0 \) or both, furthermore,

\[ \lambda_j (X_j - \bar{X}_j) = 0, \quad \lambda_j \geq 0, \quad X_j \geq \bar{X}_j. \]

Consider assets with \( X_j > 0. \) If the constraint is not binding for some \( j, \) i.e. if \( X_j > \bar{X}_j, \) then \( \lambda_j = 0, \) and it follows from (3) that

\[ E(R_j) - R^* = -\text{cov}(\frac{u'(C_1)}{E(u'(C_1))}, R_j), \]

where \( R^* \) is defined as

\[ R^* = \frac{u'(C_0)}{\theta E(u'(C_1))}. \]

\( R^* \) is similar to a marginal rate of substitution between time-zero consumption and expected time-one consumption. If the investor chooses the holding of a riskless asset freely, then
(4) is satisfied for this asset. Since the covariance will be zero, \( R^* \) will in this case be set equal to the return on the riskless asset.

Equation (4) shows that at the optimum with no binding constraint on asset \( j \), the investor will require an expected excess return (in excess over \( R^* \)) which depends on the covariance between the return and the marginal utility of time-one consumption.\(^8\) For most assets there will be a positive covariance between its return and time-one consumption.\(^9\) Since marginal utility is a decreasing function, the covariance with marginal utility will then be negative, and loosely one can say that the risk premium, the right hand side of (4), is larger the more highly the return is correlated with time-one consumption.\(^10\)

In a case with \( X_j = \bar{X}_j \), one has \( \lambda_j > 0 \), and

\[
\text{E}(R_j) - R^* = -\text{cov}(\frac{u'(C_1)}{E(u'(C_1))}, R_j) - \frac{\lambda_j}{P_j 0 E(u'(C_1))}.
\]

This completes the description of the solution to (1).

Consider now the following investment opportunity, presented to the investor after (1) has been solved. An additional unit of asset \( j \) becomes available at the price \( I \). This may e.g. be thought of as a real investment opportunity, with an output at
time one which will sell at the random price $P_{j1}$.

If the investor has chosen the holding of asset $j$ with no binding constraint, the criterion for an additional investment is of course whether $I$ is smaller than or greater than $P_{j0}$. Suppose now that the constraint has been binding. If the investor gets the investment cheap enough, he or she will still want to buy it. The effect of accepting is the same as that of increasing the constraint $\bar{\lambda}_j$ by one unit, namely a loss of $\lambda_j$ measured in utils, or $\lambda_j / u'(C_0)$ measured in real (consumption) terms. At the same time a rent of $P_{j0} - I$ is earned immediately, and the criterion is therefore to invest if the sum of these is positive, i.e. if

$$P_{j0} - I - \lambda_j / u'(C_0) \geq 0.$$ 

An underlying assumption has to be that the investor is not able to sell the investment in the market at the price $P_{j0}$ immediately. In that case the rent $P_{j0} - I$ could be reaped, and there would be no need to incur the loss $\lambda_j$. The obligation to hold on to the investment until its random return is revealed is an essential characteristic of investing in a non-marketable asset.\textsuperscript{11}

Define now the return on the investment opportunity to be $R_I = P_{j1} / I$. Plug in $\lambda_j$ from (6) and rewrite the investment
criterion (7) as

\[ E(R_I) - R^* \geq -\text{cov}\left(\frac{u'(C_1)}{E(u'(C_1))}, R_I\right). \]  

This shows that the investment criterion formally can be given exactly the same form as the criterion for investing in an "unconstrained" asset, which follows from (4). Observe, however, that \( R_I \) appears on both sides of (8). To solve explicitly for the required expected return, rewrite (8) as

\[ E(R_I) \geq \frac{E(P_{j1})}{u'(C_1)} R^*, \]

under the assumption that the denominator, the risk adjusted price, is positive. 12

The criterion (9) gives the required expected return on an investment as a risk correction factor, the large fraction, multiplied by \( R^* \), which plays the role of one plus the riskless interest rate. The risk correction factor is independent of the price \( P_{j0} \). This price, \( P_{j0} \), will only serve to determine whether a constraint is binding. Given that it is, there is a complete equivalence with the formulation in section five in which the non-marketable asset has no market price at time zero.
The risk correction factor is the expected price at time one, divided by the risk-corrected price at time one. The risk-corrected price is the expected price plus the covariance of the price with normalized marginal utility of time-one consumption. Since this covariance for most assets is negative, I have written a double minus sign in (9).

Observe that when asset $j$ is effectively constrained from below, and $\lambda_j > 0$, the following relations hold,

\[
(10) \quad E(R_I) - R^* \geq -\text{cov}(\frac{u'(C_1)}{E(u'(C_1))}, R_I) > -\text{cov}(\frac{u'(C_1)}{E(u'(C_1))}, R_j) > E(R_j) - R^*.
\]

The first inequality is the investment criterion. The relevant covariance expression for the investment exceeds that for the constrained asset $j$, which exceeds the expected excess return on asset $j$. Neither of the latter two can thus be used as the required expected excess return for this investor for the asset. In the case of no binding constraint, the last three expressions collapse to the same magnitude.

Observe that nothing in this two-period model breaks down even if all asset holdings are effectively constrained. In that case the optimization degenerates. Until the investment
opportunity comes up, the investor has no choice at all. Nevertheless he or she is able to evaluate the investment opportunity, and the relevant risk correction is a covariance measure.

3. The Multi-Period Model.

This section will show that the result of section two carries over to a multi-period model as long as consumption each period is chosen optimally, but regardless of opportunities to compose an optimal portfolio of risky assets. The attention will be restricted to investments which are committed and paid for at once, and which yield output one period later. The expression to be maximized is

\[
U = E_1( \sum_{\tau=1}^{\infty} \beta^{\tau-1} u(C_{\tau})),
\]

where \( E_1 \) is the expectation contingent on the information available at the beginning of period one.

The wealth at the end of period \( t-1 \), \( W_{t-1} \), is used for consumption in period \( t \), and for investment in \( n+1 \) assets in quantities \( X_{jt} \) at prices \( P_{jt} \). The budget constraint is

\[
W_{t-1} = C_{t} + \sum_{j=0}^{n} X_{jt}P_{jt}.
\]
As viewed from the beginning of period \( t \), the wealth at the end of that period is random, and consists of the return on the \( n+1 \) assets plus an exogenous random part, \( H_t \).

\[
W_t = \sum_{j=0}^{n} X_{jt} P_{jt}, t+1 + H_t.
\]

The return \( R_{jt} \) is now defined as \( P_{jt}, t+1 / P_{jt} \).

Using the technique of stochastic dynamic programming, define the value function at the beginning of period \( t \) as

\[
J_t(W_{t-1}) = \max_{\tau = t} E_t(\sum_{\tau = t}^{\infty} \theta^{\tau - t} u(C_{\tau})),
\]

where the maximization consists in choosing a strategy, and the expectation is contingent on that strategy. 13 By the principle of optimality,

\[
J_t(W_{t-1}) = \max E_t(u(C_t) + \theta J_{t+1}(W_t))
\]

\[
= \max_{X_{0t}, \ldots, X_{nt}} u(W_{t-1} - \sum_{j=0}^{n} X_{jt} P_{jt}) + \theta E_t(J_{t+1}(\sum_{j=0}^{n} X_{jt} P_{jt}, t+1 + H_t))
\]

s.t. \( X_{jt} \geq \bar{X}_{jt}, j = 0, \ldots, n. \)

As in section 2, the asset holdings each period are constrained from below. This gives the Lagrangian
(16) \[ L_t = u(W_{t-1} - \sum_{j=0}^{n} X_{jt}P_{jt}) + \theta E_t(J_{t+1} + \sum_{j=0}^{n} X_{jt}P_{jt}, t+1 + H_t) \]

\[ + \sum_{j=0}^{n} \lambda_{jt}(X_{jt} - \bar{X}_{jt}). \]

The first-order conditions for a maximum are, for \( j = 0, \ldots, n, \)

\[ \frac{\partial L_t}{\partial X_{jt}} = u'(C_t)(-P_{jt}) + \theta E_t(J_{t+1}'(W_t)P_{jt}, t+1) + \lambda_{jt} = 0. \]

This is very similar to (3), except that \( J_{t+1}'(W_t) \) appears instead of \( u'(C_{t+1}) \). However, by the envelope theorem,

\[ \frac{\partial J_t}{\partial W_{t-1}} = u'(C_t), \]

and similarly for \( \frac{\partial J_{t+1}}{\partial W_t} \). The latter expression can therefore be replaced by \( u'(C_{t+1}) \) in (17), and the multi-period case becomes completely analogous to the two-period case. An investment criterion like (9) can be derived.

The condition for using (9) in a multi-period model is that consumption is chosen optimally from a budget, which implies (18). This is intuitively reasonable. If instead the model had specified consumption as e.g. a fixed number, or a function of the increase in wealth (i.e. of income), there would be no reason that the covariance with marginal utility
of consumption would reflect the project return's impact on total utility (i.e. the value function) from period t+1 onwards.

For consumption to be chosen optimally, at least one asset holding must be chosen also (for the budget constraint to hold). This means that in the multi-period model, at least one asset holding must be endogenous, unconstrained, for (9) to hold. But observe that this may well be the holding of a riskless asset. The holdings of all risky assets may be effectively constrained, which means that (9) is the investment criterion even if the investor is not allowed to solve a portfolio problem by choosing a composition of risky asset holdings.

4. Financing and Reinvestment.

Whether and how an investment decision should depend on the financing of the investment, and possibly also on the reinvestment of the return, is a much debated topic. In particular in public finance the issue has not been settled. The multi-period model of the previous section can be used to analyze the question from a portfolio point of view. For a public finance application, one must be willing to postulate a social welfare function of the type (11).

The main result of this section is that one should compare the
risk-corrected return on the investment with the risk-corrected return on the financing, where risk means covariance with marginal utility of consumption. More generally, both the investment project and the financing should be shadow-priced by their impact on consumption.

By use of the envelope condition (18), the investment criterion can be simplified. It is in fact sufficient to consider the effect of an investment via consumption in the two periods affected. Observe first that (12) and (13) imply

\begin{equation}
C_t = \sum_{j=0}^{n} P_{jt} (X_{jt} - X_{j,t-1}) + H_{t-1}.
\end{equation}

Plugging this into (11), one finds

\begin{equation}
\frac{\partial u}{\partial X_{jt}} = u'(C_t)(-P_{jt}) + \theta E_t(u'(C_{t+1}P_{j,t+1}).
\end{equation}

It was shown in section three that the criterion for investing in more of asset j is whether (20) is positive or not. Consider now a situation where the investment of a quantity \(dX_{jt}\) is financed by reducing the holding of asset k. To obtain the same amount of money, one must have \(dX_{kt} = -(P_{jt}/P_{kt})dX_{jt}\). The effect on \(U\) is
\[ (21) \quad dU = \frac{\partial U}{\partial X_{jt}} dX_{jt} + \frac{\partial U}{\partial X_{kt}} dX_{kt} \]

\[ = \{ u'(C_t)(-P_{jt}) + \theta E_t(u'(C_{t+1})P_{j,t+1}) \} dX_{jt} \]

\[ + \{ u'(C_t)(-P_{kt}) + \theta E_t(u'(C_{t+1})P_{k,t+1}) \} dX_{kt}. \]

If the investor can obtain an unconstrained optimum with respect to asset \( k \), then the last line is zero, and we are back to (20). At an unconstrained optimum the marginal disutility of reducing \( C_t \) is necessarily equal (per money unit earned from the reduction) to the marginal disutility of reducing \( X_{kt} \).

If, however, the last line of (21) is not zero, rewrite (21) as

\[ (22) \quad dU = \theta P_{jt} E_t(u'(C_{t+1})) dX_{jt} \]

\[ \ast \{ [E_t(R_{jt}) - (-\text{cov}_t \frac{u'(C_{t+1})}{E_t(u'(C_{t+1}))}, R_{jt})] \}

\[ \ast \{ [E_t(R_{kt}) - (-\text{cov}_t \frac{u'(C_{t+1})}{E_t(u'(C_{t+1}))}, R_{kt})] \}. \]

The investment should be undertaken with this financing if the expression in curly brackets in the last two lines of (22) is positive. Line two gives the risk-corrected return on asset \( j \). The criterion is that this must exceed line three, the
risk corrected return on asset k.

In public finance, a frequently used argument in the
discussion about the social discount rate under full certainty
is that the appropriate rate is the producer rate of
interest,\(^{14}\) since a public investment will replace an equally
large private investment. The latter presumption is not
obvious, but this is not my point. Under uncertainty, even if
an equally large private investment is replaced, the expected
rate of return on private sector investment is not the correct
required expected rate of return.

This follows from (22) if asset k is "private sector
investment", and the public investment to be decided upon is
asset j. Only in the special case when both covariance terms
are equal, e.g. zero, will the criterion be whether
\(E_t(R_{jt}) > E_t(R_{kt})\), where \(E_t(R_{kt})\) is now the expected return on
private sector investment. In this case the question reduces
to whether the public investment yields more of the same thing
than the investment it replaces.

But in general under uncertainty, one cannot from
\(E_t(R_{jt}) > E_t(R_{kt})\) conclude that \(R_{jt} > R_{kt}\) in all states of the
world. One does not compare "the same thing". One should
instead compare the risk-adjusted returns in square brackets.
This shows that \(E_t(R_{kt})\) minus one is not in general a correct
discount rate. As a general rule, a public project should be valued directly by its impact on $U$. This is the uncertainty analogue to using the consumer interest rate as a social discount rate. But then, project financing should be shadow priced by the same principle.

This is an extension of the argument in Feldstein (1972) against the use of any other social discount rate than the consumer interest rate, but for using shadow prices on project financing.

One may want to extend the argument in Bradford (1975) and Lind (1982) about considering the reinvestment of project returns, not only financing, along the same lines. For this to be straightforward, a necessary assumption is that the reinvestment pattern is the same in all states of the world. This seems unrealistic.

5. The Mean-Variance Two-Period Model.

This section will draw the lines between the two-period model of section two and the model of Mayers (1972). Under the mean-variance assumption, the individual investment criterion can be expressed in individually observable variables, namely individual wealth. It is generally not true, however, that it can be expressed in terms of aggregate observables, as attempted by Frank and Schnabel (1984). I will therefore look
closer at the conditions for aggregation.

Since there will be aggregation across investors, a subscript $i$ is needed to denote investor $i$. His/her random wealth at time one is

\begin{equation}
W_{i1} = \sum_{j=0}^{n} X_{ij} P_{ij} + H_i + X_i P_{i1},
\end{equation}

where the exogenous random income is in two parts, one is $H_i$, the other is the investment project to be evaluated, formally considered exogenous, with quantity $X_i$ and random price $P_{i1}$.

Following Mayers (1972), I will use a utility function which is ordinal in time-zero consumption, $C_i$, the expected value of $W_{i1}$, $E_i$, and the variance of $W_{i1}$, $V_i$. The investor wishes to solve

\begin{equation}
\max_{C_i, X_{i0}, \ldots, X_{in}} u_i(C_i, E_i, V_i) \text{ s.t. } W_{i0} = \sum_{j=0}^{n} X_{ij} P_{j0} + C_i.
\end{equation}

At this stage, a cost of obtaining the random income $H_i + X_i P_{i1}$ is not explicit, but already deducted from $W_{i0}$. The Lagrangian is
(25) \[ L_i = u_i(C_i, E_i, V_i) + \lambda_i(W_{i0} - \sum_{j=0}^{n} X_{ij}P_{j0} - C_i). \]

The first-order conditions are

(26) \[ \frac{\partial L_i}{\partial C_i} = \frac{\partial u_i}{\partial C_i} - \lambda_i = 0, \]

and, for \( j=0, \ldots, n, \)

(27) \[ \frac{\partial L_i}{\partial X_{ij}} = \frac{\partial u_i}{\partial E_i} E(P_{j1}) + 2 \frac{\partial u_i}{\partial V_i} \text{cov}(P_{j1}, W_{i1}) - \frac{\partial u_i}{\partial C_i} P_{j0} = 0, \]

where \( \lambda_i = \frac{\partial u_i}{\partial C_i} \) has already been substituted in. The envelope theorem gives

(28) \[ \frac{\partial L_i}{\partial W_{i0}} = \lambda_i, \]

and

(29) \[ \frac{\partial L_i}{\partial X_I} = \frac{\partial u_i}{\partial E_i} E(P_{I1}) + 2 \frac{\partial u_i}{\partial V_i} \text{cov}(P_{I1}, W_{i1}). \]

Suppose an investment opportunity appears: The quantity \( X_I \) can now be increased by one unit by paying the investment cost \( I \) at time zero. The effect on utility of the increase in \( X_I \) is given by (29), while the cost amounts to reducing \( W_{i0} \) by \( I \), which (from (28)) has the utility effect \( \lambda_i I = (\partial u_i/\partial C_i) I \). The
investment opportunity should therefore be accepted if

\[(30) \quad \frac{\partial u_i}{\partial E_i} E(P_{i1}) + 2 \frac{\partial u_i}{\partial V_i} \text{cov}(P_{i1}, W_{i1}) - \frac{\partial u_i}{\partial C_i} \geq 0.\]

Introduce the marginal rates of substitution \(dE_i/dV_i = -(\partial u_i/\partial V_i)/(\partial u_i/\partial E_i)\) and \(dE_i/dC_i = -(\partial u_i/\partial C_i)/(\partial u_i/\partial E_i)\). Then both (27) and (30) can be summarized in

\[(31) \quad E(R_j) - 2 \frac{dE_i}{dV_i} \text{cov}(R_j, W_{i1}) + \frac{dE_i}{dC_i} \geq 0.\]

As an optimality condition, (31) holds with equality for all assets \(j\) the holding of which the investor is able to adjust freely. As an investment criterion, (31) says that the investor will want to invest in more of an asset \(j\) which fulfills (31) with inequality.

Let, for each \(i\), \(M_i\) be the set of all assets \(j\) such that if \(j \in M_i\), then (31) holds with equality for investor \(i\) for asset \(j\). Observe that (31) will also hold with equality for the return on any portfolio which is a linear combination of assets in \(M_i\). Assume that at least two of the assets in \(M_i\) are risky, with returns that have non-zero covariances with \(W_{i1}\). For all such assets,
\[
\frac{dE_i}{dV_i} = \frac{dE_i}{E(R_i) + \frac{dC_i}{\text{cov}(R_j, W_{i1})}}
\]

Now, use the other interpretation of (31) as an investment criterion for the I-asset, which is not an element of \( M_i \). Plug in for \( \frac{dE_i}{dV_i} \) from (32) and rewrite the investment criterion as, invest if

\[
E(R_i) + \frac{dE_i}{dC_i} \geq \frac{\text{cov}(R_i, W_{i1})}{\text{cov}(R_j, W_{i1})}[E(R_j) + \frac{dE_i}{dC_i}].
\]

The inequality (33) is an investment criterion for investor \( i \) for the investment in the I-asset, expressed in terms of properties of \( R_j \). Here \( R_j \) can be the return on any asset in \( M_i \), or on a portfolio of such assets, as long as the covariance in the denominator is not zero.

The criterion (33) is analogous to the investment criterion that follows from the Mayers (1972) extension of the CAPM, apart from his aggregation across investors. To obtain the aggregation, the simplest route is to assume that there exists a riskless asset, asset zero. Assume that for all \( i \), \( 0 \in M_i \). Then, from (31),

\[
\frac{dE_i}{dC_i} = -R_0.
\]
Accordingly (33) can be rewritten as, invest if

\begin{equation}
E(R_I) - R_0 \geq \frac{\text{cov}(R_I, W_{i1})}{\text{cov}(R_j, W_{i1})} [E(R_j) - R_0].
\end{equation}

When \( j, k \in M_i \), and both are risky assets, one has

\begin{equation}
\frac{\text{cov}(R_j, W_{i1})}{E(R_j) - R_0} = \frac{\text{cov}(R_k, W_{i1})}{E(R_k) - R_0}.
\end{equation}

Let \( A \) be a set of investors, and let \( M \) be a subset of the intersection of the \( M_i \)'s for all investors in \( A \),

\[ M \subseteq \bigcap_{i \in A} M_i. \]

Assume that \( M \) contains at least two risky assets. Call them \( j \) and \( k \). Then sum (36) over all \( i \) in \( A \) to obtain

\begin{equation}
\frac{\text{cov}(R_j, W_{A1})}{E(R_j) - R_0} = \frac{\text{cov}(R_k, W_{A1})}{E(R_k) - R_0},
\end{equation}

where \( W_{A1} \) is the time-one wealth of all investors in \( A \).

Rewrite \( W_{A1} \) as the sum of the return on these investors' holdings of assets that belong to \( M \), \( X_M P_{M1} \), and another term, \( H_A \). Let \( R_M \) be the return on the \( M \) portfolio, \( R_M = P_{M1}/P_{M0} \). Since \( R_M \) satisfies (37) as \( R_k \), one has
\[ \text{E}(R_j) - R_0 = \frac{\text{cov}(R_j, X_{M}P_{M1} + H_A)}{\text{cov}(R_M, X_{M}P_{M1} + H_A)} \left( \text{E}(R_M) - R_0 \right). \]

This is the Mayers (1972) extension of the CAPM, giving an equation for the required expected return for all assets \( j \) that all investors (in \( A \)) trade freely. The obvious characteristic of this result is that the valuation is the same for all investors. It has to be when all adjust freely to the same market price. The not-so-obvious characteristic is the simplicity of the aggregation, that wealth can simply be summed. The derivation here shows that it is not necessary to aggregate over all investors, and that it is not necessary to include all marketable assets in the "market portfolio" \( M \).

It is illuminating to describe the investment criterion in terms of the maximal investment cost one is willing to incur to obtain one quantity unit with price \( P_{i1} \) one period ahead. From (33) and (34), with \( R_M \) replacing \( R_j \) in (33), this criterion for a non-marketable asset is

\[ I \leq \frac{1}{R_0} \left[ \frac{\text{E}(P_{i1}) - \text{cov}(P_{i1}, W_{i1})}{\text{cov}(R_M, W_{i1})} \right] \frac{\text{E}(R_M) - R_0}{\text{cov}(R_M, W_{i1})}. \]

In the accounting literature, Boatsman and Baskin (1981), and Frank and Schnabel (1984), the valuation of a non-marketable asset has been suggested as
The difference between (39) and (40) is that the individual investor's own holding of non-marketable assets is conspicuously absent from (40). In (39), the valuation depends on the covariance of the uncertain price $P_{i1}$ with the investor's own wealth, $W_{i1}$. In (40), the valuation depends on the covariance with aggregate wealth, $W_{A1}$. This is not correct in general for imperfectly diversified investors. The composition of $W_{i1}$ will differ between such investors. When investors do not adjust freely to one market price for the asset, the individual holding should in general influence the valuation as in (39), when "valuation" is taken to mean what the investor maximally would be willing to pay for the asset.

Boatsman and Baskin (1981) are aware of this, and make the essential assumption that the return on the non-marketable asset to be valued is perfectly correlated with the return on a marketable asset. In that case the non-marketable does not affect diversification, and (40) is indeed correct.

Frank and Schnabel (1984) make no such assumption, and do indeed attempt to value an asset which has a return that is non-linearly related to the returns on other assets. Their use of the valuation formula (40) cannot give an answer to the
question of what the investor maximally would be willing to pay for the asset. The answer to this will in general differ between investors.

6. Discussion.

The results that are derived in sections two, three, and four, were based on quite weak assumptions. This was done on purpose, to show that the investment criterion based on a covariance measure is applicable under many circumstances. For normative purposes, it is conceivable that the results have practical applications. As a descriptive theory, it is a problem that the results are not in terms of observables, which makes a direct empirical test difficult.

The generality of the assumptions allows for quite strong "negative results", i.e. rejections of alternative propositions. A project return's own variance will almost never be a correct risk measure for an investor with other sources of random income. The expected rate of return on private investment is almost never an interesting social discount rate under uncertainty.

In the literature, non-marketable assets are often associated with human capital. The reasons why human capital is non-marketable have to do with absence of enforcable contracts, i.e. slavery, and with moral hazard on part of the seller: If
labor services are sold in advance, the incentive to work diminishes. An application of the results may therefore be to the valuation of investments in education and training.

Another application, which has motivated my research, is to natural resources. It is an empirical fact that resource-rich nations do not sell away claims to unextracted natural resources to a large extent. This can be explained by the same reasons, absence of enforcable contracts, and moral hazard. In this context they are called political risks, e.g. risk of breach of contract, nationalization, regulation, and/or taxation. The resources could thus be considered non-marketable from a national point of view. Required expected rates of return for oil projects in international capital markets cannot be used as social discount rates for oil projects.

In many countries national capital markets are also out of the question as sources for the expected rates of return that citizens require of natural resource projects. The reason may be poorly developed markets, or that the natural resources in effect are owned by the government, or both. As an alternative to observed rates of return in markets, a solution may be to postulate an objective function like (11). The criterion (9) gives the required expected return.
An example of a direct application is a case in which a nation has the choice whether to extract, say, 40% or 50% of the oil in place in a reservoir, by installing, say, two or three platforms. The assumption is that whichever development solution is chosen, the rest of the oil is lost for economic purposes. In this case the question is how much oil to hold in the national portfolio during the period from investment to extraction, and the investment cost is the incremental outlay required to increase extraction from 40 to 50%.

The evaluation of potential investment projects is a task to which the CAPM is frequently applied. Assume that the mean-variance assumption is accepted. A potential investment is not already traded in capital markets, but there are nevertheless cases in which it can be valued by the CAPM. In the standard CAPM world, the condition for this is that the return can be expressed as a linear combination of the returns on existing assets. Observe that in this case it is in principle not necessary to calculate the covariance between the project return and any other variable. The valuation of the project is simply a linear combination of observable prices. In practice, however, it may be more convenient to base valuation on covariances, or "project betas", which should give the same answer.

What then if there exist non-marketable assets? The condition
is now that the investment project's return can be expressed as a linear combination of the returns on marketable assets (the set $M$ in section five). Then its value is the same linear combination of prices of these assets, to which (40) applies.

But if not, the project will have different values to different investors, and (39) should be applied.\textsuperscript{27} Covariances are therefore in most cases needed to value non-marketable assets, but in principle never to value marketable assets.

7. Conclusion.

When to invest in a non-marketable asset is a well-defined problem. The investment criterion can be characterized based on quite weak assumptions.

The main result is that the appropriate risk measure is the covariance between the return on the investment and marginal utility of consumption. This is not based on the assumption that the investor only cares about the mean and the variance of the return. In a multi-period model the result holds if the utility function is time-additive, and consumption is chosen optimally. It is not necessary that the investor in any way is able to compose an optimal portfolio of risky assets.
The financing of an investment will in general have its own shadow price. The correct risk-correction again depends on the covariance with marginal utility of consumption.

The valuation of non-marketable assets will in general be specific to each individual. For a potential investment project, the valuation will be the same among a set of investors only if the project's return can be expressed as a linear combination of the returns on marketable assets that all of the investors hold in optimal amounts.
Notes.

1. In one sense, the models are used extensively to value such non-traded assets, namely, potential real investment projects. I will return to this below.

2. The market portfolio is a portfolio which consists of all risky assets in the market, each represented in proportion to its market value's share in the total market value of all assets.

3. The time-additivity is not a necessary assumption in the two-period model, but it is in the multi-period model for equation (18) to hold, as pointed out by Bergman (1985).

4. It is a matter of taste whether the model of section two is called a single-period model or a two-period model. Since consumption at time zero is included in the utility function, I prefer "two-period". The standard CAPM, and the model of Brito (1977), are single-period models.

5. Including dividends and other payouts, if any.

6. For notational simplicity, it has been assumed that all asset holdings are constrained by an $\bar{x}_j$. For application to situations with some unconstrained holdings, just make some constraints so small that they become non-binding.

7. The models of this paper may easily be extended to the case of binding constraints on short-selling, i.e. constraints with $\bar{x}_j = 0$, or even $\bar{x}_j < 0$. In order to use the standard Kuhn-Tucker formulation, in which $\lambda_j$ is not determined when $x_j = 0$,
and which does not allow negative $\bar{x}_j$, one can e.g. include an exogenous quantity of asset $j$ which the investor has to buy anyhow (as part of $H_t$ in equation (13)), and then let the investor choose an additional quantity of the same asset, restricted to be greater than or equal to some positive number.

8. The best-known model in which risk is related to consumption, is the consumption CAPM of Breeden (1979), see also Grossmann and Shiller (1982). In part of the literature, the consumption CAPM is taken to mean Breeden’s model, or at least a model in which the risk premium is determined by the covariance with consumption, not with marginal utility of consumption. LeRoy (1982) uses the term to include models in which the covariance with marginal utility appears. He observes (p. 193) that this formulation of the risk premium is very general. The present paper clarifies how general it is.

9. Time-one consumption is in the two-period model equal to time-one wealth, which is the sum of the returns on all assets. The asset return in focus is a part of this sum, and is most often positively correlated with most other asset returns in the sum.

10. Marginal utility is normalized by expected marginal utility, i.e. it is the relative deviations of marginal utility from its expected value that matters.

11. In the case where the asset holding is constrained from above, the similar characteristic is the opportunity to
hold (somewhat more of) the asset until its random return is revealed.

12. If not, the investment should clearly not be accepted.

13. The value function may have other arguments than wealth, but these (other state variables) are exogenous to the investor. In the case of quantity-constrained asset holdings, asset prices will be arguments of the value function.

14. The producer interest rate is the interest rate producers face after tax, to which they are supposed to equalize the marginal product of capital.

15. An additional subscript \(i\) on \(X_I\) and \(P_{II}\), to show that the investment project may be unique to the investor, is dropped to save notation.

16. In order to show that the model of this section is consistent with the one in section two, use a quadratic utility function, which is a special case of both models. This is left as an exercise to the reader.

17. That is, use (31) with \(R_I\) replacing \(R_j\).

18. Mayers (1973) obtains the aggregation without a riskless asset.

19. The attention is restricted to a set \(A\) of investors. The model accommodates a situation in which some assets are freely traded between some of these investors, but not between all of them. Such assets are excluded from \(M\). \(M\) can be thought of as the "market portfolio" in this restricted sense. The model even holds if some assets belong to all \(M_i\), but are
excluded from M.

20. The aggregation assumes no differential taxes. Some homogeneity in the investors' subjective probability distributions is also necessary, although it is possible to condition on a set of common information, see Grossmann and Shiller (1982).

21. Let the M portfolio consist of all assets in M, each in proportion to its share in the total holdings of M-assets by all A-investors.

22. At this point, it is interesting to observe that the consumption CAPM is hardly empirically testable even when all assets are marketable, see Cornell (1981).

23. The own variance of the investment project's return is not completely irrelevant, since it influences the covariance with marginal utility. If the project return is stochastically independent of all other returns and income elements of the investor, then the own variance accounts for the whole covariance.

24. In order to apply the investment criterion (9) to natural resources, one must take intertemporal and option aspects into consideration. But there are cases where the criterion may be used directly. In more complicated situations, the required expected return will often play a role in the valuation formula.

25. If international diversification had been perfect, there would be no connection between the physical allocation
of unextracted natural resources and the portfolio allocation of claims to these resources. This contradicts the facts.


27. This is true whether or not the project is a linear combination of other elements of \( W_{i1} \). The quantity \( X_I \) in (23) may be zero from the outset.
References.


Mayers, David (1973), "Nonmarketable Assets and the Determination of Capital Asset Prices in the Absence of a


CHAPTER 4.

COMMENTS ON THE SOLUTION

OF A DYNAMIC PORTFOLIO MODEL WITH OIL.
1. Introduction.

The uncertain value of the Norwegian petroleum wealth may be seen in a macroeconomic planning perspective. The central problem is to find an optimal extraction path, or rather, an extraction strategy, under uncertainty. In addition, the composition of other parts of the national wealth should perhaps be altered in order to get a diversified national portfolio.

The first and most widely read model dealing with these questions is Aslaksen and Bjerkholt (1985a), AB85a hereafter. The same model appears in Aslaksen and Bjerkholt (1985b) and (1986). A continuous time version, Aslaksen and Bjerkholt (1983), will not be discussed here. The present paper will discuss their model, the way they define the problem, various assumptions, and the proposed solution. I shall show that the solution is invalid due to three different mathematical mistakes. The discussion below suggests that the problem as defined by AB85a has no closed-form analytical solution.

Of their assumptions, special attention is paid to the assumed stochastic properties of the oil price, and the ability of the social planner to buy and sell oil. I show that for two by two combinations of alternative assumptions on these two points, the partial effect on utility of a change in the oil price will be positive in one case, negative in one case, and
zero in two cases.

The model represents an important attempt to use the methods of stochastic dynamic programming to solve an explicit dynamic portfolio model for the management of national wealth. The present paper will not do justice to all the virtues of this attempt. Although one should not be too optimistic about the possibilities of applying exact results from such models in the management of a national economy, I believe there are many qualitative lessons to learn. I am therefore supportive of the general way in which Aslaksen and Bjerkholt address these problems.

2. The Model.

National wealth at the beginning of period \( t, G_{t-1} \), can be used as consumption in period \( t, C_t \), or invested in different amounts in \( n+2 \) asset types. \( W_{it} \) is the value of the investment in asset \( i, i = 0, \ldots, n. \) In addition to these \( n+1 \) assets, \( S_t \) is the quantity invested in (held as) oil, to the net price \( q_t = p_t - c_t \), where \( p_t \) is the gross oil price, and \( c_t \) is the marginal extraction cost of oil.\(^2\) Below, \( c_t \) will be a function of \( S_{t-1}. \) It is assumed that the value of oil is \( q_t S_t. \)\(^3\) This gives the per-period budget constraint
(1) \[ G_{t-1} = C_t + \sum_{i=0}^{n} W_{it} + q_t S_t. \]

Next period's wealth is determined by the outcome of the stochastic return to the capital invested this period:

(2) \[ G_t = \sum_{i=0}^{n} W_{it}(1+r_{it}) + q_{t+1} S_t, \]

where \( r_{it} \) is the rate of return on \( W_{it} \). As viewed from the beginning of period \( t \), all \( r_{it} \)'s and \( q_{t+1} \) are stochastic. The return to capital is the only source of national income, and it is assumed that the composition of national wealth can be altered frictionlessly every period. In particular, oil can be sold and bought at the net price \( q_t \) for every \( t \). Together (1) and (2) describe the intertemporal budget constraint.

The stochastic assumptions are that all \( r_{0t} \) are deterministic and equal to \( r_0 = \rho_0 \), while for each \( t \), \( r_{it}, i = 1, \ldots, n \), and \( p_{t+1} \) are multinormally distributed with \( E(r_{it}) = \rho_i \) for all \( t \), \( E(p_{t+1}) = \pi_{t+1} \), and the variance-covariance matrix of

\[
\begin{bmatrix}
  r_{1t} \\
  \vdots \\
  r_{nt} \\
  p_{t+1}
\end{bmatrix}
\]

is

\[
\begin{bmatrix}
  \sigma_{11} & \cdots & \sigma_{1n} & \tau_1 \\
  \vdots & \ddots & \vdots & \vdots \\
  \vdots & \ddots & \sigma_{nn} & \tau_n \\
  \tau_1 & \cdots & \tau_n & \tau^2
\end{bmatrix}
\]

for all \( t \).

For \( t \neq s \), \( (r_{1t}, \ldots, r_{nt}, p_{t+1}) \) is independent of...
\( (r_1, \ldots, r_n, p_{s+1}). \)

Below, it will be shown that the assumption about the stochastic properties of the oil price is crucial. (The important point is not that the expected oil price differs across periods.) The return on the oil asset is \( q_{t+1}/q_t \), and this will not be independent across periods. For instance, a low outcome for the oil price in period \( t \) does not affect the expected oil price in period \( t+1 \), so that for a low \( p_t \) and any given \( c_t \) and \( c_{t+1} \), the expected return \( q_{t+1}/q_t \) becomes relatively high.

The decision maker starts out in period 1 with some wealth \( G_0 \) and a net oil price \( q_1 \). The purpose is to select a strategy in order to maximize, under the given budget constraints, an intertemporal time-additive expected utility function

\[
E_{\text{strategy}} \left[ \sum_{t=1}^{T} (1+\delta)^{-t} U(C_t) + (1+\delta)^{-T-1} V(G_T) \right],
\]

where the expectation is contingent on the strategy, \( \delta \) is a utility discount factor, \( U \) is the per-period utility function and \( V \) is the utility of final wealth. The planning period is finite, and the utility of final wealth may be thought of as representing the utility of consumption possibilities after period \( T \).
The functions \( U( \cdot ) \) and \( V( \cdot ) \) are specified as

\[
(3) \quad U(C_t) = -B \exp(-\beta C_t)
\]

\[
V(G_t) = -G \exp(-\gamma G_t),
\]

where \( B, \beta, G, \gamma \) are positive parameters.

The solution technique is that of stochastic dynamic programming. At the beginning of period \( t \), before the decision on how to spend or invest \( G_{t-1} \) is made, define \( J_t \) as the maximum value under the strategy of the expected utility of future consumption and of \( G_T \), i.e.

\[
(4) \quad J_t = \max E_t \left[ \sum_{\tau=t}^T (1+\delta)^{\tau-t} U(C_\tau) + (1+\delta)^{T-1} V(G_T) \right].
\]

By the principle of optimality,

\[
(5) \quad J_t = \max E \{ U(C_t) + J_{t+1}/(1+\delta) \}.
\]

\( J_t \) is a function of the state variables of the problem, and of all exogenous parameters. AB85a do not express clearly what the state variables are. On p. 298 they seem to suggest \( G_{t-1} \) and \( S_{t-1} \), while in the formal equations they use only \( G_{t-1} \).
However, it is clear that for the decision problem at the beginning of period $t$, the quantity of $S_{t-1}$ is in itself irrelevant, since its value is included in $G_{t-1}$, and since it can be sold without transaction costs. However, future opportunities are not determined solely by $G_{t-1}$. The investment opportunities at the beginning of period $t$ are better the lower $p_t$ is, since the expected oil price next period is independent of $p_t$. Lower $p_t$ means higher expected return on oil from $t$ to $t+1$. This is not the case for any other asset, since their rates of return are independent across periods.

In fact, the relevant oil price at the beginning of period $t$ is the net oil price $q_t = p_t - c_t$. The assumption is that the marginal extraction cost $c_t$ is a function of $S_{t-1}$, and in this way $S_{t-1}$ can be included as a state variable. However, $c_t$ can be included just as well. In either case, $p_t$ must also be included as a state variable. But in order to keep the number of state variables as low as possible, one should observe that it is sufficient to include $q_t$ as state variable in addition to $G_{t-1}$. I therefore choose as state variables $(G_{t-1}, q_t)$.

The first step towards the solution is to solve the problem at the beginning of period $T$. The problem is

$$J_T(G_{t-1}, q_t) = \Max E\{U(C_T) + V(G_T)/(1+\delta)\}$$
\[ = \text{Max} \{ U(C_T) + \frac{EV(G_T)}{(1+\delta)} \}, \]

where the maximization is with respect to \( W_0T, \ldots, W_nT, \) and \( S_T \) and subject to (1) and (2).\(^7\) AB85a use the result that when a variable \( x \) is normally distributed and \( U(x) \) is an exponential function, then

(7) \[ EU(x) = \tilde{U}(x), \]

where

(8) \[ \tilde{x} = Ex - 1/2 \alpha \text{ var}x, \]

where \( \alpha \) is minus the exponential coefficient of \( U(x) \), also equal to the coefficient \(-U''/U'\) of absolute risk aversion.

Now, \( G_T \) is indeed normally distributed. This means that with

(9) \[ \tilde{G}_T = \tilde{E}G_T - 1/2 \gamma \text{ var}G_T, \]

\[ \tilde{E}G_T = \sum_{i=0}^{n} W_i T (1 + p_i) + (n_{T+1} - c_{T+1}) S_T, \]

\[ \text{var}G_T = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} W_i T W_j T + \tau^2 S_T^2 + 2 \sum_{j=1}^{n} \tau_j W_j T S_T, \]

one has
(10) $J_T(G_{T-1}, q_T) = \max \{U(C_T) + V(G_T)/(1+\delta)\}$.

To obtain the first-order conditions, substitute for $C_T$ from (1) and for $G_T$ from (9) to get

(11) $J_T(G_{T-1}, q_T) = \max_{\{W_{iT}\}, S_T} \{U(G_{T-1} - \sum_{i=0}^{n} W_{iT} - q_T S_T) + \theta V[\sum_{i=1}^{n} W_{iT}(1+\rho_i) + (n_{T+1} - c_{T+1})S_T]$

\[ - \frac{\gamma}{2} \left( \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} W_{iT} W_{jT} + \gamma^2 S_T^2 + 2 \sum_{j=1}^{n} \tau_j W_{jT} S_T \right) \}, \]

where $\theta = 1/(1+\delta)$.

AB85a use the marginal extraction cost function $c_{t+1} = m/S_t$, implying that $\partial(c_{t+1} S_t) / \partial S_t = 0$. This gives the first-order conditions

(12) $U'(C_T) = \theta V'(G_T)(1+\rho_i - \gamma \sum_{j=1}^{n} \sigma_{ij} W_{jT} - \gamma \tau_i S_T)$,

for $i = 0, \ldots, n$, and

(13) $q_T U'(C_T) = \theta V'(G_T)(\eta_{T+1} - \gamma \tau^2 S_T^2 - \gamma \sum_{j=1}^{n} \tau_j W_{jT})$. 
These \( n+2 \) equations together with (1) for \( t = T \) determine the \( n+3 \) unknowns \( W_{0T}, \ldots, W_{nT}, S_T, \) and \( C_T. \)

Below, it will be important that the oil asset and the other assets no longer appear symmetrically in (12) – (13), as opposed to their symmetric appearance in (9). This is so because one differentiates with respect to \( W_i T \) and \( S_T \), not \( W_i T \) and \( q_T S_T \), and because of the endogenous cost.

For \( i = 0 \), (12) gives

\[
(14) \quad U'(C_T)/\theta V'(c_T) = 1 + r_0.
\]

So far, my solution is identical to the one in AB85a. The next steps are not explicit in AB85a, and the first minor divergence emerges at my equations (17) and (18) below.

Now, substitute (14) into equation (12) for \( i = 1, \ldots, n \), and into (13), to get a system of \( n+1 \) linear equations in \( W_{1T}, \ldots, W_{nT}, S_T, \) of the form \( ^8 \)

\[
(15) \quad \sum W_T = y^{-1} D_T,
\]

where the \( (n+1)\times 1 \) vectors \( W_T \) and \( D_T \) are defined by
\[ W_t = \begin{bmatrix} W_{1t} \\ \vdots \\ W_{nt} \\ S_t \end{bmatrix}, \quad D_t = \begin{bmatrix} \rho_1 & -r_0 \\ \vdots \\ \rho_n & -r_0 \\ n_{t+1} - q_t (1 + r_0) \end{bmatrix}. \]

Assuming \( \Sigma \) is non-singular, there exists a solution to (15) given by

(16) \[ W_T = y^{-1} \Sigma^{-1} D_T, \]

or

(17) \[ W_{iT} = y^{-1} \{ \sum_{j=1}^{n} (\rho_j - r_0) \hat{\sigma}_{ij} + \hat{\tau}_i (n_{T+1} - (1 + r_0) q_T) \}, \quad i=1, \ldots, n, \]

(18) \[ S_T = y^{-1} \{ \sum_{j=1}^{n} (\rho_j - r_0) \hat{\tau}_j + \hat{\tau}^2 (n_{T+1} - (1 + r_0) q_T) \}, \]

where \( \hat{\sigma}_{ij}, \hat{\tau}_i, \) and \( \hat{\tau}^2 \) are the elements of \( \Sigma^{-1} \).

This should be compared with equations (26) and (27) in AB85a with \( t = T \). Observe that \( \xi_0 = \beta / y \), while \( \xi_1 = 1 + \xi_0 / (1 + r_0) \).

The solutions are identical apart from the multiplicative constant term, which AB85a erroneously give as \( \xi_1 / \beta \) instead of as \( \xi_0 / \beta \).

Using these solutions, one can solve for \( C_T \) and \( W_{0T} \) from (1) (for \( t = T \)) and (14). Then, more serious divergences with
AB85a will appear. In an appendix it is shown that the solution for $C_T$ is

$$C_T = \left(1 + r_0\right)^{-1} \left(1 + r_0 \right) G_{T-1} + (2\gamma)^{-1} D_T^{T-1} D_T^{-m+\gamma^{-1} ln \phi},$$

where superscript $T$ on a matrix denotes a transpose, and where

$$\phi = B\beta (1+\delta)/G\gamma (1+r_0).$$

3. Errors in Aslaksen and Bjerkholt's solution.

Three different and independent errors in the solution in AB85a will be pointed out in this section. First, I will show that the solution (19) for $C_T$ is different from theirs. Second, I will show that neither their nor the correct expression for $C_T$ and $G_{T-1}$ allow the use of the certainty equivalence result (7) - (8) in the solution to the problem at the beginning of period T-1 or for any previous period. Third, I will show that the difference equation which they claim give their equation (22) as the solution, does not.

There is agreement that $C_T$ can be written as a linear function of $G_{T-1}$,

$$C_T = a_T G_{T-1} + b_T.$$
This is equation (21) in AB85a for $t = T$. There is also agreement that the coefficient $a_T$ equals $1/\xi_1$. Observe now that AB85a set the parameters $B = G = 1$. Equation (22) in AB85a gives

\[
(21) \quad b_T = \left( \frac{1+r_0}{r_0} \right)^\alpha \left[ \frac{\beta}{\gamma(1+r_0)} (\ln \frac{\beta}{\gamma} - \alpha - \ln(1+\delta)) \right] - \frac{1+r_0}{r_0} \alpha / \beta \xi_1 = \frac{\beta}{\gamma(1+r_0)} [\ln \frac{\beta}{\gamma} - \alpha - \ln(1+\delta)] / \beta \xi_1
\]

\[
= \frac{1}{\xi_1 \gamma(1+r_0)} [-\alpha + \ln \frac{\beta}{\gamma(1+\delta)}] = \frac{1}{\xi_1 \gamma(1+r_0)} [\chi - \ln \frac{1+r_0}{1+\delta}]
\]

\[
+ \ln \frac{\beta}{\gamma(1+\delta)} = \frac{1}{\xi_1 \gamma(1+r_0)} [\chi + \ln \frac{\beta}{\gamma(1+r_0)}],
\]

where AB85a have

\[
\alpha = \ln \frac{1+r_0}{1+\delta} - \chi,
\]

and

\[
\chi = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\rho_i - r_0)(\rho_j - r_0) \hat{\sigma}_{ij}
\]

\[
+ \left( \frac{q_t+1}{q_t} \right) -1-r_0 \sum_{j=1}^{n} (\rho_j - r_0) \hat{\tau}_j + \frac{1}{2} \left( \frac{q_t+1}{q_t} \right) -1-r_0 2^2 \tau^2.
\]

AB85a do not give a time index on their expressions $\alpha$ and $\chi$. 
The best guess is that in the expression for \( b_T \), the values for \( \bar{q}_{t+1} \) and \( q_t \) should be \( \bar{q}_{T+1} \) and \( q_T \).

Setting \( B = G = 1 \) in my solution (19) gives a \( b_T \) of

\[
(22) \quad b_T^* = \frac{1}{\xi_1 \gamma (1+r_0)} \left[ \chi_T^* + \ln \frac{\theta (1+\delta)}{\gamma (1+r_0)} \right]
\]

where

\[
\chi_T^* = (1/2) D_T \Sigma T - 1 D_T - m \gamma = \]

\[
\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\rho_i - r_0)(\rho_j - r_0) \sigma_{ij} + 2 \sum_{i=1}^{n} (\rho_i - r_0) [n_{T+1} - q_T (1+r_0)] \tau_i
\]

\[
+ [n_{T+1} - q_T (1+r_0)]^2 \tau^2 \}
\]

In spite of similarities, the two suggested solutions are different. A minor difference is the occurrence of a factor \((1+\delta)\) in the log term. This is explained by the ambiguity in AB85a, where the factor \((1+\delta)\) originally is used to discount \( V(G_T) \). Later, it is dropped, and this leads to its disappearance from the \( b_T \) expression in (21).

The more serious difference appears in the expression \( \chi \), where AB85a suggest a solution which is different from my \( \chi_T^* \). I cannot explain the origin of this error, but their \( \chi \) expression may result from an attempt to treat the oil asset
and the other assets symmetrically.\textsuperscript{11}

This concludes the discussion of the first error. I will now continue to the next step in the solution, the maximization problem at the beginning of period T-1. It will turn out that the use of the certainty equivalent result (7) - (8) is not applicable, and this would also have been the case if the $\chi$ expression of AB85a were correct.

First, we need an explicit expression for the $J_T(G_{T-1}, q_T)$ function. From (14) we get

\begin{equation}
\beta U(C_T) = (1+r_0)\Theta \gamma V(G_T).
\end{equation}

Combining with (10), this gives

\begin{equation}
J_T(G_{T-1}, q_T) = (1+\beta/\gamma(1+r_0))U(C_T) = \xi_1 U(C_T)
\end{equation}

\begin{equation*}
= -B\xi_1 \exp(-\beta C_T).
\end{equation*}

The maximization problem at the beginning of period T-1 is

\begin{equation}
J_{T-1}(G_{T-2}, q_{T-1}) = \max E\{U(C_{T-1}) + J_T(G_{T-1}, q_T)/(1+\delta)\}
\end{equation}

\begin{equation*}
= \max E\{U(C_{T-1}) - B\xi_1 \exp(-\beta C_T)/(1+\delta)\}.
\end{equation*}
In order to apply the certainty equivalence result to the function $J_T$, one would have to show that as viewed from the beginning of period $T-1$, $C_T$ is normally distributed. But it is not. Equation (19) gives the following expression for the optimal $C_T$ as viewed from the beginning of $T-1$:

\[(26) \quad C_T = \frac{1}{\ell} \left[ \sum_{i=0}^{n} W_{i,T-1}(1+r_{i,T-1}) + q_{T} S_{T-1} + \frac{1}{\gamma(1+r_0)}(\chi^* + 1n\phi) \right].\]

Both $q_T = p_T^{-m/S_{T-1}}$ and all $r_{i,T-1}$ are normally distributed. But $\chi^*$ contains the term $\left[ n_{T+1} - q_T(1+r_0) \right] 2^{\tau^2}$, which is the square of a normally distributed variable. Accordingly, the certainty equivalence result does not apply. Of course, it does not apply to the maximization problem in periods previous to $T-1$ either.

Considering instead the erroneous $b_T$ expression from AB85a, my equation (21), a similar problem appears. Their $\chi$ expression contains the term $\left[ q_{T+1}/q_{T-1} - r_0 \right] 2^{\tau^2}$, which contains the square of the inverse of a normally distributed variable.

However, it is possible to find an analytic expression for $E_{T-1}(J_{T-1}(G_{T-1}, q_T))$. The expression is cumbersome to derive. Unfortunately, it seems impossible to solve the first-order conditions for the maximization problem (25). Then no
analytic expression for \( J_{T-1}(G_{T-2}, q_{T-1}) \) can be found, and the problem formulated by AB85a remains unsolved.

The third serious error in the suggested solution in AB85a will just be mentioned briefly. It is of less interest since it occurs in an attempt to solve a difference equation which in itself is incorrect due to the second error pointed out above. The difference equation

\begin{equation}
\begin{aligned}
b_{t+1} = \left(\frac{\xi_{T-t+1}}{(\xi_{T-t+1}-1)}\right) b_t + \alpha / \beta
\end{aligned}
\end{equation}

is supposed to give a general solution for \( b_t \). The suggested solution (their equation (22)) is based on an erroneous terminal condition (the first error discussed in this section). But in addition, the expression \( \alpha \) is treated as a constant, while in fact, it is stochastic and varies across periods as pointed out above.

4. A CARA Model without Oil.

The second and third error which were pointed out in section 3 each seems to imply that there does not exist a tractable solution to the model. On the other hand, these problems would clearly disappear if oil were excluded from the model. This section will discuss the "remaining" model if oil is excluded.
The solution method in AB85a would then be valid. I will not give all details here. The remainder of their equation (26) is now a correct solution for $W_{it}$ for all $t$ (a generalization of my equation (17)) apart from the time index on $\xi_{T-t+1}$. The correct equation is

$$W_{it} = \frac{\xi_{T-t}}{\beta} \left\{ \sum_{j=1}^{n} (p_{ij} - r_0) \sigma_{ij} \right\}, \quad i=1, \ldots, n,$$

where

$$\xi_{T-t} = \left( \frac{1}{1+r_0} \right)^{T-t} (\beta/Y - \frac{1+r_0}{r_0}) + \frac{1+r_0}{r_0}.$$

The simplicity of equation (28) is striking: $W_{it}$ does not depend on the outcomes of the stochastic variables in periods up to $t$. In particular, $W_{it}$ is independent of $G_{t-1}$. Whatever the level of national wealth, the absolute size of the investment in any risky asset is the same. This result is known from Hakansson (1970) for a model with infinite horizon and no natural resources. AB85a do not comment on this aspect of the result.

The only variation in $W_{it}$ for a particular $i$ across different periods results from the expression $\xi_{T-t}$. This is completely deterministic. The solutions for $C_t$ and $W_{0t}$ will, however, depend on the outcomes of the stochastic process,
since they depend on $G_{t-1}$.

It is clear from (28) that the values invested in different assets will be in the same proportions across periods, i.e. for all pairs $(s,t)$ and all pairs $(i,j)$, $W_{is}/W_{js} = W_{it}/W_{jt}$. This means that the whole set of risky assets can just as well be treated as an aggregate. Only the outcomes $G_{t-1}$ affect the optimal $C_t$ and $W_{0t}$, but the process of the outcomes is described exactly by the aggregate risky asset.

Even without the assumption of a normal distribution, the CARA utility function gives independence between wealth on one hand and the values invested in risky assets on the other. I will show this in a simple two-period model: 17

Consider two individuals with the same exponential utility function, but with different levels of initial wealth, $G_i$. (Index $i$ denotes individual $i$, $i=1,2$.) Each of them may consume $C_i$ in period one, invest $W_{i0}$ in a riskless asset with rate of return $R_{0}-1$, and invest $W_{i1}, \ldots, W_{in}$ in $n$ risky assets with uncertain rates of return $R_{1}-1, \ldots, R_{n}-1$:

$$G_i = C_i + W_{i0} + \sum_{j=1}^{n} W_{ij}$$

is the budget constraint.
Their expected-utility functions are of the form

\[(30) \quad \exp(-\alpha C_i) + \theta \exp(-\alpha(W_i^0 R_0 + \sum_{j=1}^{n} W_{ij} R_j)),\]

where \(\alpha\) is a positive risk-aversion parameter, \(\theta\) is a utility discount factor, and the last parenthesis is the wealth of period two.

I shall now assume that they choose different vectors \((W_{i1}, \ldots, W_{in})\), and show that this leads to a contradiction. (It is assumed that they have equal beliefs about the probability distribution of the \(R_j\)'s, and that the individuals are not indifferent about the choices of different vectors.)

Since individual 1 could have chosen the same vector as individual 2 (while still consuming \(C_1\), but readjusting the investment in the riskless asset), but did not, he or she must prefer the chosen vector, i.e.

\[(31) \quad \exp(-\alpha C_1) + \theta \exp(-\alpha(W_{1}^0 R_0 + \sum_{j=1}^{n} W_{1j} R_j)) > \exp(-\alpha C_1) + \theta \exp(-\alpha((G_1 - C_1 - \sum_{j=1}^{n} W_{2j} R_0 + \sum_{j=1}^{n} W_{2j} R_j)).\]
Similarly, individual 2 could have chosen to invest in risky assets as did individual 1. Since another vector was chosen by 2, this gives an inequality similar to (31). Subtract now the $C_i$ terms from both sides. Observe that the exponential of the $\alpha$ times the value of a risk-free investment is deterministic, and take it out of the expectation. This gives

\[(32) \quad \exp(-\alpha W_{10} R_0) \exp(-\alpha \sum_{j=1}^{n} W_{1j} R_j) > \exp(-\alpha (G_1 - C_1 - \sum_{j=1}^{n} W_{2j} R_0) R_0) \exp(-\alpha \sum_{j=1}^{n} W_{2j} R_j),\]

and

\[(33) \quad \exp(-\alpha W_{20} R_0) \exp(-\alpha \sum_{j=1}^{n} W_{2j} R_j) > \exp(-\alpha (G_2 - C_2 - \sum_{j=1}^{n} W_{1j} R_0) R_0) \exp(-\alpha \sum_{j=1}^{n} W_{1j} R_j).\]

Multiply both sides of (33) by

\[\exp(-\alpha (G_1 - C_1 - \sum_{j=1}^{n} W_{2j} - W_{20} R_0)),\]

and use (31) to simplify the right-hand side. This gives an inequality which is identical to (32), except that the inequality signs are opposite. This is a contradiction, which completes the proof.
I have shown, both in a simple model and in the model of AB85a, that the values of the holdings of risky assets are independent of wealth when the utility function exhibits CARA. In my opinion, this property of the CARA function is counter-intuitive in the model of AB85a. A more interesting specification would be one where the $W_{it}$'s were dependent on $G_{t-1}$.

It may be possible to introduce oil into the model in such a way that there exists a tractable solution. This may make the $W_{it}$ variables dependent on other variables than time. But it is a bad property of the solution if this dependence results only from the introduction of oil, and disappears when oil is taken out of the model. The model would then show the effects of introducing oil into a rather peculiar economy.

The conclusion of this section is that the assumption of CARA gives a solution which is rather counter-intuitive. It should therefore be avoided.

5. How to Include Oil?

The previous section showed that excluding oil from the model yields a tractable, but not very appealing, solution. In this section I shall discuss some of the basic problems of including oil in the model. Unfortunately, I shall not
provide an answer to how it should be done.

One tempting solution would be to try to treat oil exactly as any other asset in the model. However, this is not possible. One reason is that oil is usually supposed to be heterogeneous with respect to extraction costs. "Oil" is really oil in the ground, unextracted, and the return $q_{t+1}/q_t$ for one unit of oil with one extraction cost is different from that of another unit of oil with another extraction cost.

One might instead consider treating oil units with different extraction costs as different assets. If the value of each unit is supposed to be its net price $p_t - c_h$, where $h$ is the index for extraction costs, the problem then is that the variance-covariance matrix of returns becomes singular, such that there do not exist solutions to the $W_{it}$ equations. The economic meaning of this is that one of the oil assets will give a better expected return than all others, whereas nothing is gained in terms of diversification from investing in many of them. The solution is therefore a corner solution, to hold only oil with the highest extraction cost. This is clearly unrealistic.

In AB85a, the value of the holding of oil from the beginning of period $t$ to the beginning of period $t+1$ is $q_t S_t =$
$(p_t - c_t(S_{t-1}))S_t$, where $c_t$ is the marginal extraction cost.

This assumption is not well explained. Under full certainty, it is standard to assume that the value of one unit of unextracted oil is $p_t - c_h$. But the way this is extended to the value of a whole set of oil units under uncertainty entails several problems.

The first problem is that one would expect $S_t$ to consist of units of oil with different extraction costs. In that case, the value should also reflect this. A more natural assumption would be that the value is $p_t S_t - \int c(S) dS$. For a marginal adjustment this might not be important, but the wealth effect of $S_t$, on e.g. consumption, would certainly be different. The other possibility is that $S_t$ consists of units of oil with the same extraction cost, $c_t$. This is very unrealistic, particularly with the assumption that $c_t$ is a function of $S_{t-1}$.

It is not unusual to assume that marginal extraction cost is a decreasing function of $S$. This is often done in models where the owner of the oil extracts and sells oil from an initial stock, such that $S$ is a decreasing function of time. In AB85a, however, $S$ may both increase and decrease from one period to the next. Accordingly, the assumption is that when the decision maker buys oil, these units of oil satisfy the same $c_t(S_{t-1})$ function. This is not explained, and seems to
be counter-intuitive.

All trade in oil at the beginning of period \( t \) takes place at the same price \( q_t = p_t - c_t(S_{t-1}) \). Even if oil is only sold, not bought, one would expect this to be a good approximation only when few units of oil are sold. At the beginning of period \( t \), when oil is sold away to reduce the holding from \( S_{t-1} \) to \( S_t \), one would expect to encounter oil units with progressively higher \( c_t \)'s as more oil is sold.

Under uncertainty, I believe the correct valuation can only be found by taking the option value into account. Owning a unit of unextracted oil has an option value: One does not have to extract the oil. The valuation of oil options is discussed e.g. in Tourinho (1979), Pindyck (1981), Paddock, Siegel and Smith (1984), and Brennan and Schwartz (1985). An obvious example on the importance of the option value aspect is the case where the oil price in some period is below the extraction cost \( c_h \) for some unit of oil. The value of this unit may still be strictly positive if there is a chance that the oil price will exceed \( c_h \) in some later period, provided that unextracted oil has low or zero storage cost.

For most stochastic specifications, there will be some such chance. With the stochastic specification of AB85a, the chance is definitely there. The normal distribution of \( q_t \)
does not rule out any outcome, and if oil prices are very low one period, this does not affect the distribution of future oil prices.

This leads to a discussion of the plausibility of the stochastic specification of the oil price process. In their numerical application, AB85a assume that the expected oil price grows exponentially, which means that their assumption can be formulated as

\[ p_t = p_0 \exp(\mu t) + \epsilon_t, \]

where \( \mu \) is a deterministic exponential growth rate and the \( \epsilon_t \)'s are stochastic, i.i.d. The most common alternative is

\[ p_t = (1+r_{xt})p_{t-1}, \]

where \( r_{xt} \)'s are stochastic and independent across time, perhaps i.i.d. Of course, (34) and (35) do not comprise an exhaustive list of alternatives. But they are useful, simple formulae for illustrative purposes.

From the theoretical point of view, in a model where oil producers can turn extraction on and off every period (as the price becomes known), (34) seems to be an unrealistic stochastic process for equilibrium oil prices. A low outcome
of $p_t$ would imply that the expected return on oil over the next period is very high, and this would induce producers to hold back extraction. Similarly, a high outcome of $p_t$ would induce high extraction. This means that the supply of oil would fluctuate very much between periods. If in addition oil could be bought and sold without transaction or storage costs, oil would represent a very good investment in periods with low $p_t$'s and vice versa. This would reinforce the fluctuations.\footnote{19}

A market equilibrium price path under these simplifying assumptions is more likely to be one where the expected return on oil from one period to the next is more or less stable. This suggests (35).

On the other hand, oil is not actually bought and sold for investment purposes to a large extent. The consumption demand and factor input demand for oil are more likely to fluctuate around a stable expected path, which suggests (34). These sources of demand constitute a major difference between commodities, like oil, and financial assets, a difference which is not captured by extending the Hotelling-type models to a world of uncertainty. Moreover, many oil producers resist to turn extraction on and off every period because of high fixed costs or because of revenue targets. Also, it may take time to turn extraction on or off, which means that the oil price may have changed in the meantime. These are some
reasons why the arguments of the two preceding paragraph may not be valid.

In fact, during the past thirteen years oil prices have risen sharply, but also fallen during some periods. Some observers have maintained all the time that the expected future oil prices follow an exponential path starting at the oil price at that time. This is consistent with (35), but in retrospect it does not seem very realistic. Ex post it is tempting to say e.g. that the expected return on oil before the first OPEC shock was very high, while it was very low after the second OPEC shock. This is much closer to (34).

In order to explore the consequences of the two types of oil price processes, I shall discuss two different models informally. The models are of the stochastic dynamic programming type. Let model A be like the one in AB85a, i.e. a model which allows the decision maker to adjust the holding of oil up or down every period without transaction costs. Let model B be an opposite extreme, a model in which the oil extraction in each future period is exogenous, independent of realizations of oil prices.

This defines four cases, A-(34), A-(35), B-(34), and B-(35). In all cases, the wealth inherited from the previous period, $G_{t-1}$ is an argument in the $J_t$ function. It turns out that the
$J_t$ function will have the oil price as an argument in only two of these cases, and that its first partial derivative with respect to the oil price is negative in one of these, positive in the other!

The case A-(34) is found in AB85a, and I have discussed in section 2 why the oil price enters as an argument in the $J_t$ function. The higher the oil price, for a given value of wealth, the worse are the investment possibilities, since it is more expensive to invest in oil. That is, $\partial J_t(G_{t-1}, p_t)/\partial p_t$ is negative.

However, in the case A-(35) the investment opportunities are not affected by the current oil price, and $p_t$ will not be an argument in the $J_t$ function.

The case B-(34) also excludes $p_t$ from the $J_t$ function. Remember that the effect of $p_t$ on the value of current extraction is included in the value of wealth, $G_{t-1}$. Future extraction values will not be affected by the current oil price.

In the case B-(35), however, a high current oil price has positive implications for the expectations about future oil prices, and therefore the value of future extraction. This gives a positive $\partial J_t(G_{t-1}, p_t)/\partial p_t$. 
In order to summarize, I can make a four-field table of the sign of $aJ_t/\partial p_t$, table 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>(34)</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>(35)</td>
<td>0</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 1: The sign of the first partial derivative of the $J_t$ function with respect to the oil price in two different models under two different assumptions about the oil price process.

Of course, the partial derivative does not bring out the total effect of a change in the oil price, since such a change will also affect wealth. But the results of e.g. a high oil price are rather different in the four cases, and this shows that the distinctions between these models and oil price processes are important. In discussions of consequences of oil price uncertainty in applied work, the underlying assumptions are not always made clear. This may confuse the discussions.
Notes.

1. In Aslaksen and Bjerkholt (1985b) there are a few additional typing errors and errors due to changes in notation.

2. AB85a use the symbol $b_t$ for marginal extraction cost. However, they use the same symbol for something else, and I reserve "$b_t$" for that use, see equation (20) below.

3. For the moment, I will not discuss the assumption that the value of oil is quantity times (price minus marginal extraction cost).

4. Aslaksen and Bjerkholt are aware that this assumption implies that the stock of oil is both increased and reduced from time to time, and that this seems unrealistic for many oil producers today.

5. In their numerical application, AB85a use data from Norwegian national accounts for the returns on different types of capital. For those data, the asset returns are hardly independent across periods.

6. The terminology in AB85a is slightly misleading when it is stated that they assume "stochastic independence between rates of return, including the oil price" (p. 298).

7. AB85a drop the factor $1/(1+\delta)$ as they go from the first to the last line of (6). In the sequel, they most places drop it, and of course, for the last period's problem, it can be included in the constant $G$.

8. The intermediate steps are
\[
l_1 r_0 = 1 + \rho_1 - \gamma (\sum_{j=1}^{n} \sigma_{ij} w_j T + \tau_i S_T), \quad i = 1, \ldots, n
\]

\[
l_1 + r_0 = q_T^{-1} (n_{T+1} - \gamma (\sum_{j=1}^{n} \tau_j w_j T + \tau^2 S_T)).
\]

This gives

\[
\sigma_{11} w_{1T} + \cdots + \sigma_{1n} w_{nT} + \tau_1 S_T = \gamma^{-1} (\rho_1 - r_0),
\]

\[
\vdots
\]

\[
\sigma_{nn} w_{1T} + \cdots + \sigma_{nn} w_{nT} + \tau_n S_T = \gamma^{-1} (\rho_n - r_0),
\]

\[
\tau_1 w_{1T} + \cdots + \tau_n w_{nT} + \tau^2 S_T = \gamma^{-1} (n_{T+1} - q_T (1 + r_0)).
\]

9. The peculiarities of the results (17) and (18) will be discussed in section 4 below.

10. The expression $\xi_{T-t}$ is defined on pages 301 - 303 in AB85a. Given the way AB85a formulate the maximization problem in period $t$, it is obvious that the correct expression in their equations (26) and (27) should be $\xi_{T-t}$.

11. In fact, their $\chi$ expression does not treat all assets symmetrically, since the elements of $\Sigma^{-1}$ appear symmetrically. The matrix $\Sigma$ is not the covariance matrix of $(r_{lt}, \ldots, r_{nt}, q_{t+1}/q_t)$, which would have made it symmetric in all asset returns, but of $(r_{lt}, \ldots, r_{nt}, q_{t+1})$. 
12. In fact, there is another difference equation in $\xi_{T-t}$ which makes the solution to (27) tractable as long as $\alpha$ is treated as a constant.

13. If (27) had been the correct specification, with $\alpha$ constant, the solution in AB85a would indeed have been a solution to the difference equation.

14. Since $W_0$ is allowed to be negative, nothing prevents the investment of all the $W_{it}$'s.

15. In fact, AB85a write "Given the optimal consumption, the accumulation in the uncertain assets is determined as a one-period portfolio problem," and then they give their equations for $W_{it}$ and $S_{it}$, which also are independent of $G_{t-1}$, and of $C_t$. The phrase "Given the optimal consumption" is misleading.

16. When $\beta/\gamma < (1+r_0)/r_0$, as is the case in the numerical application in AB85a, $\xi_{T-t}$ decreases in $t$. If $T$ is very large, $\xi_T$ is close to $(1+r_0)/r_0$, then $\xi_{T-t}$ decreases and reaches the value $\xi_0 = \beta/\gamma$.

17. Thanks to Aanund Hylland for suggesting this proof.

18. This is e.g. the discrete-time version of the assumption in Pindyck (1981). It is also most common in the option valuation literature, and therefore convenient if one wants to obtain explicit option value formulae. This is of course no decisive argument for using (35)!

19. In Aslaksen and Bjerkholt (1985b), the variation in the holding of oil across periods is surprisingly low.
References:


Appendix.

In this appendix, I will show how equation (19) is derived from equations (1) (for \( t = T \)), (14), (17) and (18). (14) gives

\[(19a) \quad B \beta \exp(-\beta C_T) = \theta (1+r_0) G \gamma \exp(-\gamma G_T).\]

Taking logs and using \( \phi = B \beta (1+\delta)/G \gamma (1+r_0) \) gives

\[(19b) \quad \ln \phi - \beta C_T = -\gamma G_T.\]

For all \( t \), let \( \bar{q}_t = n_t - c_t \), and define the \((n+1)x1\) vectors

\[
R_t = \begin{bmatrix} 1+r_1 \\ \vdots \\ 1+r_n \\ \bar{q}_t \end{bmatrix}, \quad R_0 = \begin{bmatrix} 1+r_0 \\ \vdots \\ 1+r_0 \\ \bar{q}_t(1+r_0) \end{bmatrix}, \quad \Delta_t = \begin{bmatrix} \rho_1 - r_0 \\ \vdots \\ \rho_n - r_0 \\ n_t + 1 - 2c_t + 1 - q_t(1+r_0) \end{bmatrix}.
\]

Then,

\[
EG_t - W_0 (1+r_0) = W_T^T R_t + 1,
\]

\[\text{var} G_t = W_T^T \Sigma W_t.\]

From (19b) one gets
(19c) \[ C_T = \frac{\gamma}{\beta}(W_0T(1+r_0) + \frac{W^T}{T R_T + 1} - (\gamma/2)W^T_S W_T) + \beta^{-1}ln\phi \]

\[ = \frac{\gamma}{\beta}(1+r_0)(G_{T-1} - \sum_{i=1}^{n} w_i T q_i S_T) \]

\[ + (\gamma/\beta)(W^T_{T R_T + 1} - (\gamma/2)W^T_S W_T) + \beta^{-1}ln\phi, \]

which gives

(19d) \[ C_T (1+\gamma(1+r_0)/\beta) = \beta^{-1}ln\phi \]

\[ + (\gamma/\beta)(1+r_0)(G_{T-1} - \sum_{i=1}^{n} w_i T q_i S_T) + \frac{W^T}{T R_T + 1} - (\gamma/2)W^T_S W_T \]

\[ = \beta^{-1}ln\phi \]

\[ + (\gamma(1+r_0)/\beta)G_{T-1} + (\gamma/\beta)\left[-\frac{W^T}{T R_0 T + \frac{W^T}{T R_T + 1} - (\gamma/2)W^T_S W_T} \right] \]

Using (16) and symbols previously defined, this gives

(19e) \[ C_T = \frac{1}{\xi_1}G_{T-1} + \frac{1}{\gamma(1+r_0)}D_T \Sigma^{-1}(-R_0 T + R_{T+1} \frac{1}{2} \Sigma^{-1} D_T) \]

\[ + \frac{1}{\gamma(1+r_0)}ln\phi. \]

Now, the three vector terms in parentheses can be rewritten as
(19f) \(-R_{0T} + R_{T+1} - (1/2) \Sigma^{-1} D_T = -R_{0T} + R_{T+1} - (1/2) D_T = (1/2) \Delta_T\). 

Premultiplying this by \(2D_T^{T \Sigma^{-1}}\) gives 

(19g) \[ D_T^{T \Sigma^{-1}} \Delta_T \]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} (\rho_i - r_0)(\rho_j - r_0) \hat{a}_{ij} + \sum_{i=1}^{n} (\rho_i - r_0)(\eta_{T+1} - q_T(1+r_0))
\]

\[-2c_{T+1} \hat{\tau}_i + \sum_{i=1}^{n} [\eta_{T+1} - q_T(1+r_0)] \hat{\tau}_i (\rho_i - r_0)
\]

\[+ [\eta_{T+1} - q_T(1+r_0)]^2 \hat{\tau}_i \]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} (\rho_i - r_0)(\rho_j - r_0) \hat{a}_{ij} + \sum_{i=1}^{n} (\rho_i - r_0)[\eta_{T+1} - q_T(1+r_0)] \hat{\tau}_i
\]

\[+ [\eta_{T+1} - q_T(1+r_0)]^2 \hat{\tau}_i \]

\[-2c_{T+1} \sum_{i=1}^{n} (\rho_i - r_0) \hat{\tau}_i + [\eta_{T+1} - q_T(1+r_0)]^2 \hat{\tau}_i \).
\]

Of the last three lines, the first two are equal to \(D_T^{T \Sigma^{-1}} D_T\). 

From (18), the very last line is equal to \(-2c_{T+1} \gamma S_T = -2m \gamma\). 

(19) follows from (19e) since 

(19h) \[ \gamma^{-1} D_T^{T \Sigma^{-1}} (-R_{0T} + R_{T+1} - (1/2) \Sigma^{-1} D_T) = (2\gamma)^{-1} D_T^{T \Sigma^{-1}} \Delta_T \]

\[= (2\gamma)^{-1} (D_T^{T \Sigma^{-1}} D_T - 2m \gamma) = (2\gamma)^{-1} D_T^{T \Sigma^{-1}} D_T - m. \]
CHAPTER 5.

LESS THAN SINGLE DIVIDEND TAXATION.
1. Introduction.¹

The purpose of this study is to describe an existing special form of taxation of corporate profits which gives high-income individuals an incentive to incorporate. This will be contrasted to the description of corporate profits taxation in King (1977), where this kind of incentive is assumed to be non-existing because of its adverse effects.

It is well known that double taxation of dividends makes dividend income unattractive under many tax regimes. Single taxation would mean equality between dividends and other cash income with respect to taxation. What will be described here, however, may be regarded as a case of "less than single taxation" of dividends.

King (1977, section 3.2) defines a general way of evaluating the taxation of corporate profits, when distributed as dividends, in terms of two parameters. He then develops four different categories of corporate profits tax systems. This note will describe a fifth category which is a necessary generalization of one of King's categories in order to describe the dividend taxation in Norway. A factual error in King (1977) will be corrected.

Capital gains taxation is neglected in this note as in the relevant sections of King (1977). Since capital gains are
taxed even less heavily than dividends, their existence will make it even more profitable to incorporate, if anything.

Much of the discussion of the relation between personal and corporate taxes has concentrated on capital gains taxation, which in most tax systems is very lenient. However, this does not lead to a general incorporation of all sources of income, because individuals may have a need for regularly receiving cash, and because capital gains are difficult to realize, especially in small, non-traded companies. This is the reason for considering dividend taxation separately.

2. Notation.

The following notation is from King (1977):

\[ Y = \text{pre-tax profits}, \]
\[ D = \text{dividends in the hands of the shareholders, net of all taxes}, \]
\[ T = \text{total tax liability on profits (firm's plus shareholders' liability)}, \]
\[ m = \text{personal marginal tax rate}, \]
\[ \tau = \text{rate of corporate income tax on undistributed profits}, \]
\[ \theta = \text{the opportunity cost of retained earnings in terms of net dividends foregone}, \]

where \( \tau \) and \( \theta \) are defined by eq. (3.2) in King (1977):
(1) \[ T = \tau Y + \left[ \frac{(1-\theta)}{\theta} \right] D. \]

Since capital gains taxation is neglected, the tax liability, \( T \), is simply the basic rate, \( \tau \), times the pre-tax profits, \( Y \), plus what additional taxes follow from the distribution of dividends. When retained earnings are traded for dividends, for each unit reduction of retained (post-tax) earnings, \( 1-\theta \) is taxed away, while the shareholders receive (net of all taxes) \( \theta \). \( D \) increases by \( \theta \), and those additional taxes are therefore \( (1-\theta)(D/\theta) \). In general, \( (1-\theta) \) represents combined effects of taxes on the corporate and personal level, and it may even be negative.

King (1977, pp. 56-57) compares the taxation of corporate income with the taxation of unincorporated income. For this purpose he introduces his ACID (Attempted Corporate Integration of Dividend Taxation) test statistic. Observe that of an extra unit of pre-tax corporate profits, shareholders will receive \( \theta(1-\tau) \), while of an extra unit of pre-tax unincorporated profits, they will receive \( (1-m) \). The ratio of these two numbers is the ACID test statistic.
3. The Two-rate System.

One of the four tax system categories analyzed by King taxes distributed profits at a lower rate, \( c_d \), than the rate on undistributed profits \( c_u \). On the personal level, dividends are taxed at the rate \( m \), so that gross dividends, \( G \), can be defined by

\[
(2) \quad D = (1-m)G.
\]

The total tax burden is then

\[
(3) \quad T = c_u(Y-G) + c_dG + mG = c_uY + (m+c_d-c_u)G,
\]

and from (1) we have

\[
(4) \quad \tau = c_u
\]

\[
\theta = (1-m)/(1+c_d-c_u).
\]

Also,

\[
(5) \quad ACID = (1-\tau)/(1+c_d-\tau).
\]

Since both \( \tau \) and \( c_d \) have values between zero and one, ACID will also be in that interval. King (1977) assumes that all governments will want ACID \( \leq 1 \), since ACID > 1 makes it
profitable to incorporate to avoid taxes. Governments presumably want to prevent this.

4. The 2x2-rate System.

There exist tax systems which do not fit any of King's four categories. They can, however, be described by his parameters. Let me define the 2x2-rate system as a generalization of the two-rate system. It has two rates both at the personal and at the corporate level. At the personal level, dividends are taxed at the marginal rate \( m_d \), while other income is still taxed at the marginal rate \( m \).

Redefine now \( G = D/(1-m_d) \) as gross dividends, and use this instead of \( G \) in (3). Then (3) and (4) will hold for the 2x2-rate system if \( m \) is replaced by \( m_d \). Specifically, we get

\[
(6) \quad \theta = (1-m_d)/(1+c_d-c_u).
\]

But the rate \( m \) is still needed to compare the taxation of corporate profits to that of other income. The ACID becomes:

\[
(7) \quad \text{ACID} = \frac{\theta(1-\tau)}{(1-m)} = \frac{(1-m_d)(1-\tau)}{(1+c_d-\tau)(1-m)},
\]

and we can no longer be certain that it is below unity. Also, it will depend on personal tax rates unless \( (1-m_d)/(1-m) \) is
equal for all shareholders.

5. The Norwegian Tax System.

Since 1969, Norway has had a 2x2-rate system, with the additional relation

\[(8) \quad m = c_d + m_d.\]

A few explanations may be of interest. We have:

\[m_d = \text{the shareholder's (progressive) national personal marginal income tax rate},\]
\[m - m_d = \text{the shareholder's (proportional) municipal personal income tax rate}.\]

The rate \(m\) is the sum of a national and a municipal rate. The municipal rate has been 0.23 for the last 15 years, and cannot therefore be neglected. At the personal level, distributed dividends are deductible only against the municipal tax.

At the corporate level, \(p\) is also a sum of a national and a municipal rate. Here, distributed dividends are deductible only against the national tax. The municipal rate is the same for individuals and corporations, and appears as \(c_d\) in (8). The total tax burden on the corporate and personal levels becomes:
\[(9) \quad T = c_u (Y - \tilde{G}) + c_d \tilde{G} + m_d \tilde{G}.\]

From (1) we get

\[(10) \quad \tau = c_u', \quad \theta = (1 + c_d - m)/(1 + c_d - c_u).\]

The ACID for the Norwegian system becomes:

\[(11) \quad ACID = \frac{(1 + c_d - m)(1 - \tau)}{(1 + c_d - \tau)(1 - m)},\]

and, with \(c_d > 0\), we get

\[(12) \quad ACID \geq \frac{\tau}{\langle \tau \rangle}, \text{ as } m \geq \frac{\tau}{\langle \tau \rangle} \]

In Norway, \(\tau = 0.508\) in 1985, and a substantial proportion of shareholders will have \(m > 0.508\).

In King (1977, p. 55 - 57) the Norwegian system (as of 1975) is described as a two-rate system with \(c_d = 0\), so that ACID = 1. This is a correct description of the tax system at the national level only. There, dividends are completely deductible against corporate income tax, so that \(c_d\) is zero,
as assumed by King (1977). But because of the substantial municipal rate and the deduction rules, the Norwegian system is not adequately described if the municipal level is neglected.

To compare the effective tax rates on corporate and unincorporated profits, observe that the effective tax rate on corporate profits is

\[
1 - \theta(1-\tau) = 1 - \frac{(1+c_d-m)(1-\tau)}{1+c_d-\tau} = 0.1619 + 0.6814 m.
\]

To carry the argument to the extreme, one may assume that persons with higher marginal rates than \( \tau \) will channel their income through corporations. A person with marginal tax rate \( m_1 > \tau \) will observe that, at the margin, it pays to transfer income (if possible) so that it is earned as distributed dividends instead as of some other kind of personal income (to be taxed at \( m_1 \)).

Such a transfer would reduce the base for the progressive tax \( m_d \), since only what becomes the gross dividend part, \( (1-\tau)/(1+c_d-\tau) \) of the transferred income, will be included in the new base. Reducing the base for \( m_d \) means reducing the rate \( m_d \), so that new transfers will reduce the tax burden
until one of two things happens: either the total marginal rate $m = c_d + m_d$ is reduced from $m_1$ to below $\tau$, at which point no more transfer is profitable; or all income is incorporated, and $m$ is still greater than $\tau$.

This leaves us with a testable hypothesis: Everyone with $m > \tau$ receives no personal income directly, but instead as dividends and/or capital gains. This is clearly false in Norway. One explanation may be restrictions on the distribution of dividends. Corporations are required to build a "reserve fund". However, this is not a satisfactory explanation, since the fund is not lost. Another explanation may be the effect of wealth taxes, which have been neglected so far. These apply both at the personal and the corporate level, and formally this is double taxation. For companies that are not publicly traded, however, the wealth assessment is explicitly lenient at the personal level, and incorporation into such a company therefore seems beneficial for people in high tax brackets.

A tentative conclusion about the Norwegian system is that there exist incentives in the tax code for people in high tax brackets to channel income through corporations that are non-traded. For lower tax brackets, there is the opposite incentive. It follows that some income will be able to avoid the high statutory marginal personal tax rates. Such an
avoidance may incur extra costs, e.g. transactions costs, and it may reduce tax revenue. It may also be perceived as unfair by persons who are restricted from avoiding high marginal tax rates.

6. Discussion.

The reserve fund requirement and the wealth taxes make it difficult to draw conclusions about the Norwegian system without a very detailed analysis. The more general lesson to be learned from this study is methodological.

When studying different taxes applied to the same tax base, it is important, but not quite straightforward, to sort out how the taxes are related to each other. This is where the designers of the Norwegian scheme should have been more careful. To understand how the Norwegian system was conceived, observe that when \( c_d = m - m_d \), eq. (9) becomes:

\[
T = \tau(Y - \bar{G}) + (m - m_d)\bar{G} + m_d\bar{G},
\]

where the second term is the explicit tax on gross dividends at the corporate level, while the third term is the tax on gross dividends at the private level.

One is tempted to believe that the intention of the system was that the total tax on distributed corporate profits should be
m. If \( Y = \tilde{G} \), this seems to be consistent with (14), since we get \( T = m\tilde{G} = mY \). But this is an illusion for the following reason.

Suppose the corporation gets an additional unit of pre-tax profits. It wants to keep retained earnings at the same level as before, but will increase \( \tilde{G} \). However, if the whole unit of \( \Delta Y \) is distributed as \( \Delta \tilde{G} \), nothing is left to pay the tax \( c_d \Delta \tilde{G} \). Clearly, if retained earnings are unchanged, some of the extra \( Y \) must be channeled into the tax \( c_d \Delta \tilde{G} \). But since the change in \( Y-\tilde{G} \) is now positive, a tax \( \tau \) has to be paid on \( \Delta(Y-\tilde{G}) \). At this point the firm's tax liability is \( c_d \Delta \tilde{G} + \tau c_d \Delta \tilde{G} \). But the second term here must also be part of \( \Delta(Y-\tilde{G}) \), and this results in the additional tax \( \tau^2 c_d \Delta \tilde{G} \), which results in \( \tau^3 c_d \Delta \tilde{G} \), and so on.

This infinite geometric series has the limit \( [c_d/(1-\tau)]\Delta \tilde{G} \), which is the tax burden on \( \Delta Y \) on the corporate level. This can alternatively be seen from the following equation:

\[
(15) \quad \Delta Y = \tau(\Delta Y-\Delta \tilde{G}) + c_d \Delta \tilde{G} + \Delta \tilde{G},
\]

which implies:

\[
(16) \quad \Delta Y = [1+c_d/(1-\tau)]\Delta \tilde{G}.
\]
Eq. (15) is clearly necessary when \( \Delta Y \) is used only for \( \Delta \tilde{G} \) and the additional tax burden. From (16) we see that the tax paid on the corporate level on \( \Delta Y \) is not \( c_d \Delta Y \), it is not even \( c_d \Delta \tilde{G} \) (which still would have made the total tax rate different from \( m \)), but it is

\[
[c_d/(1-\tau)] \Delta \tilde{G} = [c_d/(1+c_d-\tau)] \Delta Y.
\]

Adding the personal liability \( \eta_d \Delta \tilde{G} \) gives the total liability \( [(\eta-\eta_d)/(1+c_d-\tau)] \Delta Y \), which we already have in (13), since

\[
1 - \frac{(1+c_d-\eta)(1-\tau)}{1+c_d-\eta} = \frac{\eta-\eta_d}{1+c_d-\eta},
\]
Notes.

1. This chapter has been published as Lund (1986).

2. In addition to the Norwegian system, the Swedish system has two rates at the personal level, as indicated by King and Fullerton (1984, pp. 112f).

3. The reserve fund must accumulate by at least 10 percent of profits until it reaches 20 percent of book value of equity. In relation to my topic there are two opposing effects of the reserve fund requirement. Accumulation of retained earnings is subject to the tax rate, \( \tau \), and when \( m > \tau \), this accumulation is beneficial compared to direct individual saving. However, the reserve fund is freely distributed only upon liquidation, and in that event the company will normally not have enough tax liability to be able to deduct this liquidation dividend completely. For a fuller explication of the reserve fund rules, see Lund (1987).

4. The wealth is taxed at a rate \( w_p \) at the personal level, and at a rate \( w_c \) at the corporate level. However, for a company which is not publicly traded, the assessment rate at the personal level, \( \phi \), is approximately 0.6. This gives a modified

\[
ACID = \frac{(1+c_u-m)(1-c_u-w_p+\phi)/r)(1-c_u)(1-m-w_p/r)}{(1+c_u-c_u)(1-m-w_p/r)},
\]
where \( r \) is the interest or capitalization rate. For Norway in 1985, \( w_c = 0.004 \), while \( w_p \) is progressive, between 0 and 0.023. Even with \( m = c_u \), the modified ACID exceeds unity when wealth is large and in the form of shares in a company which is not publicly traded.
References.


