Taxation and the effect on extraction rates[1]

by

Aksel Hiorth[2] and Petter Osmundsen[3]

University of Stavanger

Abstract

Petroleum tax design analysis usually abstracts from the fact that taxation affects not only project selection but also concept design and drainage strategy, and thereby the extraction and recovery rate. The implicit assumption is that oil companies are unresponsive to tax changes where concept selection and production decisions are concerned. In his novel approach, Smith (2014) develops a broader model which accounts for the additional effects. This provides a consistent framework which permits analysis of specific settings. Berg et al (2018) apply the model to establish the effect of a reduction in uplift on the Norwegian continental shelf (NCS) and find, for this case, that the traditional assumption of an unresponsive company still holds.

Analysing field data from the NCS, we find that the actual production profiles do not resemble the mathematically tractable theoretical model of a constant decline in production used in Smith (2014) and Berg et al (2018), and we explain why this is the case for offshore fields. We propose a novel theoretical resource model which fits the data from the 41 largest field on the NCS. It is a general resource model that allows for injection from start of production, and where the model of a constant decline in production is a special case. Our model is also a mathematically tractable model of constant decline, but where the decline is at the level of the individual well rather than the field. The recovery rate is thus responsive to the number of wells, and tax design affects the recovery rate. Our findings suggest that companies are responsive, and that tax design should take account of distortions in the selection of development concepts.

1. Introduction

In his seminal paper, Smith (2014) expands tax analysis in several dimensions in what he refers to as a parsimonious tax model. Whereas traditional tax theory only analyses the decision to invest in an exogenously given project, this model also accounts for the fact that tax design affects exploration, dimensioning of development projects, production profile and overall

---

We would like to thank Kjell Agnar Dragvik, Magne Emhjellen, Tor Fjæran, Kjell Løvås, Lars Myklatun, Svein Skjæveland, Erik Søndenå and Anders Toft for useful comments and suggestions.

[2] University of Stavanger, NO-4036, Stavanger, Norway, +47 51831757, aksel.hiorth@uis.no

[3] University of Stavanger, NO-4036, Stavanger, Norway, +47 51831568, petter.osmundsens@uis.no

production. The theoretical model of Smith (2014) accounts for the investor's simultaneous consideration of petroleum exploration, development and production. The paper opens up for a wide range of important empirical applications. Our paper is, to the best of our knowledge, the first of these.

Our empirical findings also indicate a potential for augmenting the theoretical model in Smith (2014). The latter starts by presenting a general model of taxation, a novel framework with which we agree. He thereafter applies this through the use of a specific resource model. At this point, he moves from the general to the specific. This is where we believe the model could be augmented by recognising that one specific reservoir model does not apply to all reservoirs or all extraction philosophies. The unique openness of the Norwegian petroleum directorate, that make production curves and information on production and injection wells available for all fields on their web page, make it possible to make empirical tests for offshore fields.

Where the primary production phase is concerned, Smith (2014) applies the exponential decline model of oil field development which is common in the literature. See Adelman and Paddock (1980). Reference is made to Smith and Paddock (1984), who test the cost implications of the exponential decline model for global onshore and offshore fields. The findings concur with expectations for onshore fields but contradict a priori expectations for offshore fields. Nor do our empirical findings for Norwegian offshore fields concur with the exponential decline model. This could indicate that the parsimonious tax model should differentiate between onshore and offshore fields, applying different resource models. Another factor which may call for model differentiation between offshore and onshore production is the considerably higher capital intensity in offshore production and front-end loading of investments, which calls for shorter production periods. It is also important to account for differences in operating costs. These increase over time for offshore fields and thereby affect the decision on abandonment as well as the production profile.

The distinction between resource models is not derived from location as such, but from the extraction method chosen. In the exponential decline model, it is presumed that oil companies use the natural energy in the reservoir alone to produce oil. Smith (2014) augments this model by allowing for enhanced oil recovery (EOR) after a third of the reserves have been extracted. This does not concur with our experience from the NCS or with the knowledge of the reference group we have put together for this research. The latter consists of key subsurface specialists and experts in dimensioning petroleum development projects. Its members work in Equinor (which operates the majority of the fields on the NCS and also has a large global portfolio of offshore fields), Petoro, and the Norwegian Petroleum Directorate (NPD). They point out that usual practice in the offshore sector is to drill additional water and natural gas injectors from the start of production in order to provide pressure support. This is open information, included in the data made public by the NPD.[4] According to our industry panel, this extraction policy is not specific to Norway but also followed in other offshore oil provinces. However, oil provinces vary in this respect. According to the industry panel, it is not usually possible to extract as much as a third of the oil in place from NCS fields merely by using the natural energy in the reservoir as presumed in the general model in Smith (2014). Once reservoir pressure becomes too low, moreover, it would not be economic to extract more, simply because there is not enough energy left in the reservoir to lift the fluids from the underground to the surface. To obtain a high recovery rate, it is therefore crucial to maintain reservoir pressure from the start through injection. The number of wells is decisive for the recovery rate which will be achieved.

[4] See the NPD fact pages; http://factpages.npd.no/factpages/Default.aspx?culture=no

Extraction and recovery rates are determined not only by when and how many injection wells are sanctioned, but also by the number of production wells drilled.

A crucial issue is whether the number of offshore production and injection wells primarily accelerates oil extraction or increases it (and thus the recovery rate). According to our industry panel, the former is only the case in an ideal situation with a homogeneous reservoir featuring very high permeability. The panel indicates that this is rare for offshore fields globally, and even rarer now because the most profitable fields in many oil provinces have been developed. Reservoirs are often segmented and require a large number of production wells, and injection of water and/or natural gas to maintain reservoir pressure so that a profitable proportion of the resources can be produced. Reservoirs are typically segmented owing to faults and zones of tight strata. This results in pressure differences which present extraction challenges.

According to our panel, injection wells in most cases do not primarily help to accelerate production but are necessary to extract the economic resource potential. Segmented reservoirs need injection into every segment to drain the oil. Moreover, wells in the initial business case are often not sufficient, infill wells may be needed to reach economic production targets. At this point in production, every well is a separate business case and insufficient tax incentives will mean suboptimal production.

Technical and economic conditions differ greatly between these two categories of fields. Another assumption in Smith (2014) which is not adequate for offshore fields is the simplification of constant operating costs over time. The general framework thereby opens up for new and more realistic tax studies able to account for crucial issues which have previously been ignored, and where policy recommendations will be contingent on the type of project analysed.

We examine the effect of a 2013 increase in effective taxation (by reduced uplift) on the NCS. This has been analysed by Berg et al (2018) applying the Smith (2014) framework, including the general theoretical reservoir model. An empirical analysis we have conducted for the largest petroleum fields off Norway finds that Norwegian petroleum fields do not fit the theoretical model. We discuss the effect of the tax increase when accounting for the actual production pattern. Conclusions change when we recognise the effect which a reduced number of production wells has on the recovery rate. Smith (2014) finds, by using a theoretical production curve of constant relative decline, that the assumption of unresponsive companies is valid. Using a bell-shaped production curve based on empirical analysis from the NCS, we find that tax designers need to account for responsive companies. The analytical framework developed by Smith (2014) is therefore important.

In research on tax design in the petroleum sector, the emphasis is on the decision to sanction a project – whether to invest. See Keen and McPherson (2010), for example. The analysis is often conducted in a simplified setting, where investment projects are treated like a black box. The taxation system is assumed to have no effect on the design of projects. In other words, capital investment, the production period, and extraction and recovery rates are presumed to be unaffected by a tax change. According to our industry panel, this is not realistic. Smith (2013) finds that the assumption of unresponsive companies is common in tax analyses, despite recommendations that company response be modelled in the form of endogenous design (Poterba, 2010). Our contributions are to analyse endogenous project design and petroleum tax design for offshore projects, with the emphasis on extraction and recovery rates.

Few research articles analyse endogenous design and thereby the economics of extraction and recovery rates. Osmundsen (2013) examines the effect which the choice of development concept has on the recovery rate in offshore fields, but does not address taxation. Introducing a parsimonious theoretical reservoir model, Smith (2014) addresses the issue of tax neutrality. Drawing on Smith (2014), which is based on a theoretical reservoir model appropriate for onshore fields, Berg et al (2018) analyse the production effect of the 2013 uplift reduction for offshore fields on the NCS and concludes that the implications of the resulting investment reductions are primarily to postpone production. Initial production is unchanged and only a slight production decline occurs in the final production stage. The authors conclude that the present tax research practice of addressing project sanctioning alone is a good approximation, since the effect on production and extraction rate is negligible. The primary effect is to delay production somewhat. Their results therefore support the assumption of unresponsive companies in traditional tax theory.

This addresses an interesting policy question of a more general nature. When designing policy, simplifying assumptions are typically made about company response. The latter is to a large extent a black box in existing theory. Could simplified and imprecise modelling of corporate behaviour inflict a welfare cost, in that company response differs from simplified assumptions? Berg et al (2018) conclude that this is not the case where Norwegian petroleum taxation is concerned.

Smith (2014) provides a general framework which is potentially very useful for policy design. When evaluating policy design, using updated empirical data which actually fit the current case is crucial. In this paper, we show that Berg et al (2018) fail at this point. They make no effort to use available empirical production data from the NCS. Instead, they utilise the general theoretical reservoir model used in Smith (2014) and the cost calculations in Smith (2014) based on input from Smith and Paddock (1984). Descriptive statistics from Norwegian fields quickly reveal that the declining production function assumed in Berg et al (2018) does not fit NCS. We argue that the theoretical reservoir model from onshore production does not fit current offshore fields, such as those off Norway, and that empirically realistic assumptions need to be made. We develop an alternative theoretical reservoir model for offshore fields which fits data from the largest petroleum fields on the NCS. Extraction and recovery rates are endogenous, including in the initial stage, and depend on the number of extraction and injection wells. Tax implications are discussed. The assumption of unresponsive companies is shown to be invalid in this case.

The government needs to take account of the reality that deviating from a neutral tax position distorts not only decisions on project selection but also those on concept design and the number of wells, and thereby affects the recovery rate in sanctioned projects.

2. Norwegian petroleum taxation

We will address the effect of tax design on extraction and recovery rates by analysing a case on the NCS. Norway used to have a stable net petroleum income tax system, with a special tax and an uplift. Marginal tax was set at 78 per cent and the uplift was set so that the tax system was neutral (non-distorting) for a company using the net present value (NPV) method and a nominal discount rate of around nine per cent. The uplift was 7.5 per cent over four years, to adjust for the fact that tax depreciation was over six years instead of being an immediate deduction as with a neutral cash flow tax. The role of the uplift was to shield the normal rate of return from the special tax and thus avoid underinvestment. In 2013, the uplift was reduced to 5.5 per cent, a tax shock which increased the effective tax rate defined as the share of the NPV captured by the government. The argument was that oil companies should use a partial cash flow method and apply a very low discount rate for tax depreciation since the latter is claimed to be risk free.

The problem with this approach is that decision criteria in private companies are not set by the government. Oil companies use the traditional NPV method, and the result is that projects which are profitable to society will not be sanctioned (Osmundsen et al, 2015). The question posed in this paper is how the projects which are still sanctioned will be affected by the tax increase.

In this paper, we open the box and look at decisions affecting the recovery rate in offshore projects. We acknowledge that corporate taxes may affect not only the decision to sanction the project but also project design and thereby the recovery rate. The effect of the uplift reduction has been examined by Berg et al (2018), using the theoretical framework in Smith (2014). They argue, as do Osmundsen et al (2015), that it is difficult to determine the correct cost of capital for each cash flow component. Instead, they use a traditional NPV method and a 10 per cent discount rate in their base case.

Failing to recognise a constant decline in production in NCS field data, we have examined actual production curves on the NCS using data from the NPD.

3. Reservoir model for offshore fields

To the left in Figure 1 a plot of the decline model of Smith (2014) is shown, and to the right there is a typical production curve from a field on the NCS. The curve to the right is based on a visual inspection of production curves for the 45 largest fields on the NCS.

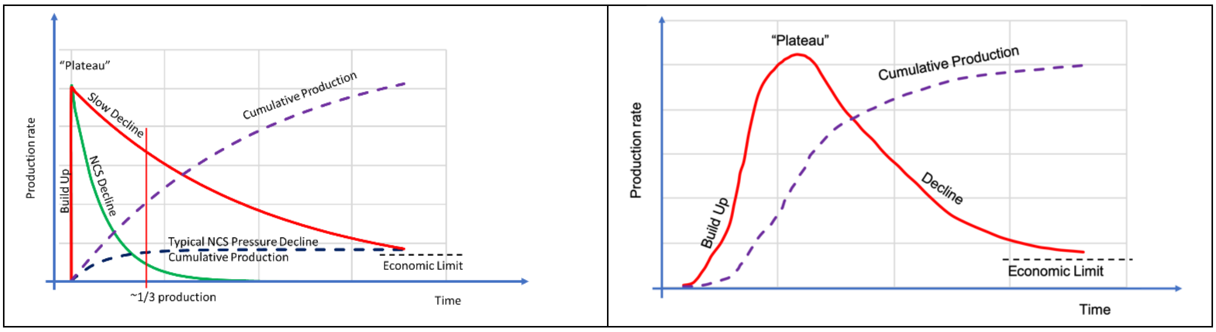

Figure 1: The red lines are the production rates and the dotted purple lines are cumulative production. Left: the production decline model used in Smith (2014). Right: A typical production curve from a field on the NCS.

Smith (2014) states that they only use the decline model for primary production, which is assumed to recover 1/3 of the original oil in place. The production rates in the theoretical decline model is quite different from actual production curves from the NCS. The reason for this is that the model of Smith (2014) is based on a different recovery strategy than what is normally used on the NCS. The decline model to the left in Figure 1 assumes that the hydrocarbons are recovered using the natural energy (fluid pressure) in the reservoir. Extraction of hydrocarbons reduces the fluid pressure, and thereby the driving force for hydrocarbon expulsion. The rate of decline and the rate of extraction are merely a consequence of the number of extraction wells drilled (Smith 2014).
The process can be visualised by thinking of a car tyre filled with gas. The gas inside the tyre has a higher pressure than the surrounding and will flow out of the tyre when we cause a puncture (i.e. drill a well). If we create more puncture, we recover the gas inside the tyre faster, but the ultimate recovery will be similar regardless of the number of holes (wells) in the tyre. So why does the production rate on the NCS differ from this decline model? If we produce oil by simply letting the

pressure decline until the production stops, the ultimate recovery is expected to be 10-20% (Arps 1956), which is indicated by the green line to the left in Figure 1. On the NCS almost all the reservoirs are produced by water injection. This is probably partly due to the great success story of water injection on the Ekofisk field. The Ekofisk field is Norway's first and second largest field, and it was planned to be produced by pressure depletion. The original estimated recovery factor was 18% (Hermansen 1997), but after a successful water injection pilot in 1987 (Sylte 1988) it was decided to extend the waterflood. Now the recovery factor is estimated to be 51% (NPD 2019), which is about $180 \cdot 10^6$ Sm^3 additional oil recovered by water injection. The main reason for the weak performance by merely applying pressure depletion is that oil (in contrast to gas) does not have a great ability to expand when pressure is decreased. The tyre example in this case is similar to an oil filled tyre; now we only recover a small fraction of the oil by puncturing the tyre. However, it is possible to recover much more of the oil by injecting a cheaper fluid that can replace the high-value oil. Thus, in order to get a high recovery, one cannot only rely on the natural energy in the reservoir to produce oil, but additional pressure support in terms of e.g. water injectors is needed. A further complication is that the reservoir often consists of isolated compartments (e.g. created by faults and fractures), thus the reservoir is not a single tyre but a series of disconnected or poorly connected tyres. The only way to produce a structurally complex reservoir is to drill more wells to reach the isolated compartments. Thus, the reservoir model should be extended to account for ways to add more energy to the reservoir in form of pressure support (e.g. water injection) to capture the field production when this is the dominant recovery mechanism. The model should also account for production wells that are drilled to reach compartments in the reservoir with limited contact with the main reservoir.

The traditional way to model fluid movement in an oil reservoir is to solve mathematical equations which describe the conservation of mass and momentum. Modern computers may take days to complete a simulation of an oil field's production history, and reservoir engineers make decisions about field development based on these simulations. The reservoir models are complicated to solve, because the equations are solved on a fine spatial grid to determine exactly where the oil and water are flowing in the sub-surface. This is important because decisions about the placement of new wells have to be made with very good knowledge about the location of the remaining oil. However, it has been recognized by several authors that models have become too complicated, because they tend to include all possible physical effects regardless of the use, and therefore impractical for many tasks. Lately researchers have argued that in some cases reduced physics models can serve an important role (Ibrahima 2017, Albertoni 2003). We support this notion, because a model should be developed with a specific purpose in mind.

In this paper, we are not interested in the spatial location of oil and water for a specific reservoir. For tax design purposes we need a simplified, representative model of what drives overall field recovery over time. Our model is described in detail in the appendix, and we only repeat the main points here. We develop a novel simulation model that borrows methodology primarily developed to investigate the transport of chemicals in flow reactors (Danckwerts 1953). Our idea is to view the fraction of oil in the reservoir as a *concentration*. This is an untraditional way of modelling fluid flow in a reservoir, but as we will demonstrate very fruitful. Water is injected (a zero concentration of oil), and a fraction of oil is produced. If the reservoir is viewed as a continuously stirred tank, oil production from it will follow an exponential decline curve, where the time constant is simply the injection rate divided by the reservoir pore volume. This result is consistent with the empirical results described in (Arps 1945), and with a more formal derivation presented by (Fetkovich 1980) using a different approach than what we present. In the Appendix, we show that the pressure decline model, applied for primary production in Smith (2014), can be extended to include the effect of water injection, and the result is:

$$Q_{\mathrm{hc}}^{field}(\mathrm{t}) = \mathrm{q} \cdot \mathrm{e}^{-\alpha \mathrm{t}} \sum_{i=1}^{N(t)} \left(1 - \Gamma\left(n, \frac{N(t) n(t - T_i)}{\tau}\right)\right) H(t - T_i), \tag{1}$$

where q is the average production rate of an individual well due to water injection and pressure decline, and α is the decline rate due to the natural energy in the reservoir. The parameter n can be viewed as a description of reservoir heterogeneities, a low number gives early water breakthrough and a high number gives a homogeneous piston displacement. $N(t)$ is the number of production wells at any given time, $H(t - T_i)$ is the Heaviside step function ensuring that a well does not contribute to production before it is put online, T_i is the time when the well is drilled, and $\Gamma(\cdot)$, is the incomplete gamma function. τ is a time constant which is related to the *total* volume of hydrocarbons in the reservoir, and the average well flow rate, q:

$$\tau = \frac{V_{hc}}{q}. \tag{2}$$

Thus τ/N will be the recovery time for a field. If we set $n = 1$, equation (1) reduces to a very tractable format:

$$Q_{\mathrm{hc}}^{field}(\mathrm{t}) = \mathrm{q} \cdot \mathrm{e}^{-\alpha \mathrm{t}} \sum_{i=1}^{N(t)} e^{-\frac{N(t)}{\tau}(t - T_i)} H(t - T_i)\,. \tag{3}$$

If there is no water injection $\tau \to \infty$, and the model reduces to a pure pressure decline model. In the following we will put $\alpha = 0$, and ignore the contribution from pressure decline. Note that our model then only has two parameters, τ and q. This model is tested against data available in the NPD fact pages (NPD 2019), where hydrocarbon production history and well information are available. During the lifetime of a producing oil field oil wells are drilled at different intervals. On the NPD website, data are available which describe when a well was drilled on a specific field and what type of well it was (e.g. oil producer or water injector). The model is implemented in Python and the "curve_fit()" function of SciPy is used to determine τ and q in equation (1). In Figure 2, four fields are shown. On the left axis the production data of oil (green), oil equivalents (red), gas (black), and our model fit (blue) is shown. The right axis shows the cumulative number of wells over time (black dashed line). The Oseberg field has $400 \cdot 10^6$ Sm^3 oil in place, and ~150 wells, the Gullfaks field has $380 \cdot 10^6$ Sm^3 oil in place, and ~150 wells, the Snorre field is smaller with $277 \cdot 10^6$ Sm^3 oil in place, and ~60 wells, and Embla is the smallest with $12 \cdot 10^6$ Sm^3 oil in place, and 7 wells. Figure 2 shows that our model is able to capture the recovery rates reasonable well, independent of field size. In the Appendix in Figure 5 a scatter plot is shown for all the fields, where we compare the total recovery predicted for each field and the actual recovery, and most of the fields are matched within 10%. It is quite remarkable that a model with so few parameters are able to fit such a wide range of fields. Most likely a perfect fit could be obtained by letting the decay rate, $\tau \to \tau_i$, i.e. give each well an independent decay rate. However, this would introduce one additional parameter for each well, making the model less transparent and much harder to fit and possibly introduce a non-unique solution.

Another interesting feature of the model is that the fitted values correlate very well with the expected values. This is shown to the left in Figure 5 in the Appendix, where we compare the product $q \cdot \tau$ with the original oil in place estimated by the companies (NPD 2019), and find that it scales linearly with the oil in place.

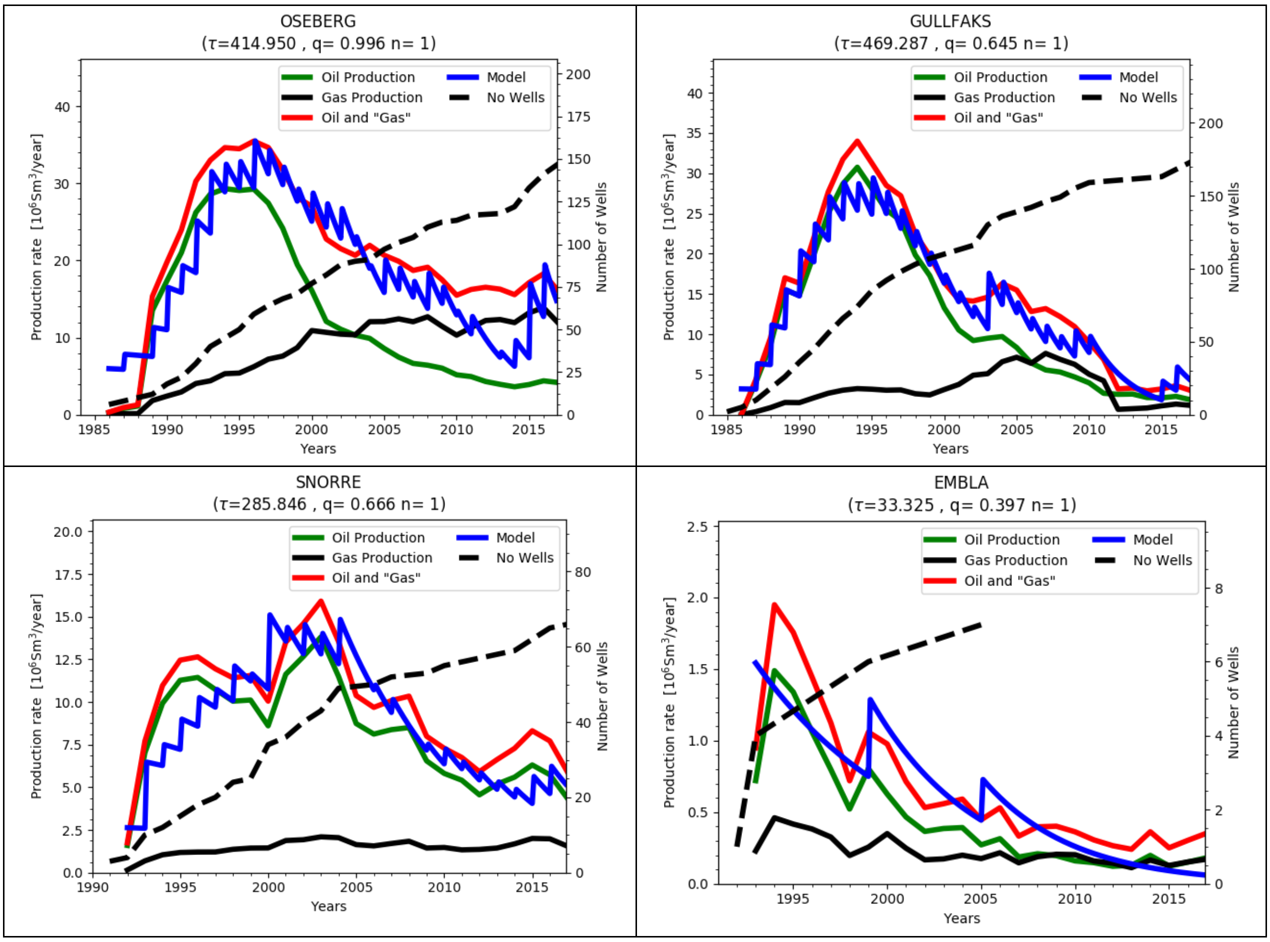

Figure 2: Modelled and actual historical production curves from the Oseberg, Gullfaks, Snorre and Embla fields. The dotted line is the cumulative number of wells drilled.

Perhaps the strongest conclusion which can be drawn from the model is that the primary driver for oil recovery is the number of wells drilled. Another interesting technical detail is that the exponential decline model (n=1) gives overall best fit to the data, thus indicating that reservoirs on the NCS are heterogeneous and that the wells will experience fairly rapid water break-through.

4. Offshore versus onshore fields

Offshore fields differ from those onshore in many respects. First, they are typically much more capital intensive, and most of the investment is often taken up front. Expanding extraction capacity is typically much more expensive once the production facility has been installed offshore. Undertaking construction or modifications offshore is several times more expensive than doing it before the production facility leaves shore, so most EOR measures need to be in place from the start. As a minimum, the facility needs load-bearing capacity and deck space for future installation of EOR equipment, implying high up-front investment. If such spending is not made initially, the EOR measures will typically never be implemented This differs from onshore production, where space is often unlimited and where a sequential approach can be selected. An offshore tax regime which leads to underinvestment is therefore likely to have less EOR capacity installed and a lower recovery rate.

Compared with an onshore field, an offshore field features much more infrastructure and much higher initial investment. To satisfy the high rates of return required from front-end-loaded investments, a major part of production must come early. The average pay-back time on the

NCS is six years.[5] After the initial build-up, production reaches the designed plateau level and, according to our industry panel, the oil companies try to maintain this for as long as possible.

Production and injection wells in midwater and deepwater areas, such as the NCS, are much more expensive than on onshore fields. Fewer wells are therefore chosen than would be typical on land. A reduction in the number of wells therefore has a more severe impact on the recovery rate. This is not accounted for in Berg et al (2018).

The decision to shut down a field also differs between onshore and offshore. Owing to the corrosive environment found offshore, maintenance costs rise over time. Reducing costs when activity has declined is also much harder in offshore operations. Since the size of the production facility is given, maintenance costs are not reduced by downscaling. It also takes more or less the same number of personnel to staff the facility regardless of whether oil or water is being produced. Our industry panel estimates that about 80 per cent of operating costs are independent of the level of oil production. During the field's late life, the proportion (cut) of water in the wellstream rises and adds to processing costs. Operating costs are therefore convex over time and not constant as assumed in Smith (2014). Processing capacity is given and cannot be extended at a later date, as would be possible onshore. When a growing part of the fixed capacity is devoted to processing water, oil production declines. According to our industry panel, rising costs and reduced production capacity often entail a sharp growth in unit cost.

5. The decision on concept design

The overarching framework in Smith (2014) opens up for addressing not only the decision on whether to develop a given field but also the choice of development concept. A next step could be to endogenise the companies' investment and concept design decisions in the parsimonious tax model.

How an investment decision is taken when using the NPV and a given requirement for rate of return is well known. However, we have been unable to find anything about the decision on concept design, including dimensioning of the development project, in academic literature. In the following, we draw on inputs from our industry panel to describe how this decision is made.

When deciding how to develop a specific petroleum field, oil companies start with a lean design which has a low cost and recovery rate. Subsequently, they check whether it would be profitable to choose a bigger concept which implies higher capital expenditure but a higher recovery rate as well. This is examined by applying a delta analysis. The cash flows of the lean and bigger development concepts are compared and the differential (delta) cash flow calculated by simply deducting the cash flow of the lean design from that of the bigger design. The NPV of the delta cash flow is then calculated. If this is positive, a larger concept is considered profitable. An even larger concept is then tested, and this process continues until an increase in scale is no longer profitable.

A distortive tax – in the form of reducing uplift below the level which secures neutrality, for example – has the effect of reducing the scale of the selected development concept, including the number of wells. That implies a loss in terms of a suboptimal recovery rate.

[5] Report from the Norwegian Climate Risk Commission, NOU 2018:17, Appendix 5.

The choice of concept design also involves decisions on the flexibility or robustness of the development solution. The concept design is determined with incomplete knowledge of the size and properties of the reservoir. When production drilling begins, surprises of varying degrees are likely to be encountered. The reservoir model could prove incomplete. If the production design adopted is flexible, adjustments can be made to partly or fully recoup the planned recovery rate. However, flexibility comes at a cost. If a low-cost design with little flexibility is selected – owing to reduced investment incentives following a reduction in uplift, for example – the planned recovery rate cannot be achieved and suboptimal production will be the result. A lower capacity and flexibility typically also imply a reduced ability to tie in other nearby petroleum deposits. Tie-in fields help to prolong production on the main field after its own output goes off plateau. These are typically marginal resources which cannot justify a stand-alone process facility, and thus depend for development on using process capacity provided by a nearby field. The consequence of suboptimal capacity and flexibility may thereby include social costs in terms of a shorter production period for the main facility and stranded resources in the vicinity.

6. Investment decision applying a social required rate of return

Berg et al (2018) refer to the fact that the government's required rate of return (the social cost of capital) is typically lower than the capital cost for the companies. This is not addressed in conventional tax analysis, which implicitly assumes that government and companies have the same required rate of return. Neutrality is defined in tax literature as a tax system which does not distort the decisions of the companies. If the decisions of the companies do not accord with the social optimum, however, typically in terms of underinvestment, the government would actually like to distort it by introducing stronger investment incentives. The overarching tax model of Smith (2014) makes it possible to account for the differing cost of capital for various dimensions of field development.

Berg et al (2018) abstract from the decision to invest and address the choice of development concept for sanctioned projects. They argue that fewer wells resulting from a reduction in uplift would be beneficial to government since the postponement of revenue is less costly to society than to the companies owing to the difference in required rates of return. In reaching this conclusion, they assume that the recovery rate will remain close to unchanged, in spite of lower investment, and that operating costs are constant over time – ie, they abstract from all the downside of reduced investment. As we have demonstrated, this downside is considerable. A leaner and less flexible development concept would lead to a lower expected recovery rate. According to our industry panel, a lower required rate of return for sanctioned projects would call for a bigger and more flexible development concept. Accounting for the fact that oil companies have a required rate of return several times higher than the government's therefore calls for stronger investment incentives (a higher uplift).

7. Conclusions

Applying the theoretical model in Smith (2014), Berg et al (2018) analyse the consequence of a tax increase on the NCS. They conclude that a reduction in the number of wells had almost no effect on the recovery rate, with only the timing of production changed. This accords with traditional tax theory which presumes unresponsive companies. Implicit assumptions for this result are homogeneous reservoirs with a very high permeability which allow the same level of drainage with fewer wells. Our alternative model, which is tested to fit empirical data from the NCS, shows the opposite: extensive drilling is crucial to the recovery rate and a tax system which deviates from neutrality will generate suboptimal production. The physical or reservoir-technical reason for this is that most reservoirs are heterogeneous, meaning that they are fractured and compartmentalised. Water injected in one well will only contact a small fraction of the reservoir and, when water breaks through in a producer, its cut will increase while oil production drops. The challenge is then to divert the water to different places in the reservoir, and to contact unflooded areas in order to displace oil. This takes a large number of wells. A high level of recovery from such reservoirs is therefore a direct result of activity and efforts by the operators. As more wells are drilled, and more production data are gathered, additional information is gained about the reservoir. This in turns affects the optimal placement of wells. A reduction in tax incentives which lowers the extent of drilling therefore has a clear and negative effect on the recovery rate. This must be taken into account in an optimal tax design. A traditional trade-off thereby also exists between revenue and tax distortions when it comes to the decisions on concept design and drilling.

Projects with a high rate of return before tax may be sanctioned in spite of a distortive tax regime, only marginal projects are affected when it comes to sanctioning. For all projects, however, decisions on project design and infill drilling will be affected by distortive corporate taxation, because these decisions in their nature are marginal. The consequence is suboptimal recovery rates. This is a problem that for the most part is ignored by the tax literature, presuming unresponsive companies.

The tax system may produce two types of distortions: 1) it may distort the decision to invest, in that projects which are profitable before tax may not be sanctioned, and 2) it may distort the design of the project and thus the production profile and recovery rate. Berg et al (2018) abstract from 1). This is problematic since they analyse a specific case, the Norwegian 2013 uplift reduction, and should cover all effects. Given the basic investment model assumption they make, which is that companies use the traditional NPV method and a 10 per cent nominal required rate of return, there are projects profitable for society which will not be sanctioned (Osmundsen et al (2015)). The 2013 tax increase was followed by a steep reduction in sanctioned investment projects on the NCS, owing partly to cost inflation related to the business cycle and partly to the tax increase. Both had the effect of reducing the breakeven price below the required level set by the companies. According to the theoretical model in Berg et al (2018), this does not happen. One obvious reason is that they assume a price of USD 100 per barrel, which is unrealistic. Other factors are that they abstract from capital rationing and that their calculations of breakeven prices are wrong. The breakeven prices specified are so low that all projects would also be sanctioned after the tax increase. This is not what we observe. Using project data from Rystad Energy, Berg et al calculate breakeven prices in their Table 5 which are much lower than Rystad Energy lists in its own publications. In their calculation, Berg et al seem to have discounted the cost (the numerator of the breakeven price) but not the production volume (the denominator), giving inaccurately low breakeven prices. Making correct calculations of breakeven prices and comparing these with observed sanctioning criteria by oil companies which ration capital (Osmundsen et al (2017)), we see that many projects are not sanctioned. On the basis of project calculations, we find that oil companies need a before-tax return above 15.4 per cent after the tax increase to generate an after-tax return of 10 per cent. They would need an even higher rate of return if they are not in a tax paying position, yet another topic not addressed by Berg et al.

In developing an alternative reservoir model for the NCS we have made many simplifications and compromises. The reservoir model we develop will most likely be inadequate for some fields. It is not meant as a model that fits all fields, but a model that makes a fair representation of a representative field. The implicit assumption is that government would like one general tax system, not a field specific tax system curtailed to individual reservoirs. If government would like a more specific tax system, e.g., to differentiate between old and new fields or oil and gas fields, more specific reservoir models are called for.

A topic for future research is to make cost calculations for fields on the NCS and develop a parsimonious resource model along the lines of Smith (2014). Such a model, that portrays the complex trade-offs presented by different tax instruments, could prove useful for policy design. Another potential application of the framework provided by Smith (2014) is to analyse unconventional oil production, in particular the expanding US production of tight oil.

References

Albertoni, A., and Lake, L. W. (2003). "Inferring Interwell Connectivity Only From Well-Rate Fluctuations in Waterfloods" *SPE* doi:10.2118/83381-PA

Arps, J. J. (1945), "Analysis of Decline Curves", *Transactions of the AIME*, 160(01): 228-247.

Arps, J. J. (1956), "Estimation of Primary Oil Reserves", Society of Petroleum Engineers, SPE-627-G.

Berg, M, Bøhren, Ø. and Vassnes, E. (2018), "Modeling the response to exogenous shocks: The capital uplift rate in petroleum taxation", *Energy Economics* 69: 442-455.

Danckwerts, P. V. (1953). "Continuous flow systems: Distribution of residence times." *Chemical Engineering Science* 2(1): 1-13.

Daniel, P., Keen, M. and McPherson, C, eds (2010), *The Taxation of Petroleum and Minerals: Principles, Problems and Practice*, Routledge, jointly with the International Monetary Fund, April, 2010, ISBN: 978-0-41578-138-1: 425-444.

Fetkovich, M. J. (1980), "Decline Curve Analysis Using Type Curves", *Journal of Petroleum Technology*, 32(06): 1065-1077.

Hermansen, H.; Thomas, L.K.; Sylte, J.E.; Aasboe, B.T. (1997), "Twenty Five Years of Ekofisk Reservoir Management", SPE-38927-MS,SPE

Ibrahima, F., Maqui, A., Negreira, A. S., et al. (2017). "Reduced-Physics Modeling and Optimization of Mature Waterfloods". *SPE* doi:10.2118/188313-MS

Levenspiel, O. 1972. Chemical Reaction Engineering, 2nd edition, New York: John Wiley and Sons

NPD, 2019. http://factpages.npd.no/factpages/

Osmundsen, P. (2013), "Choice of Development Concept – Platform or Subsea Solution? Implications for the Recovery Factor", *Oil & Gas Facilities*, Society of Petroleum Engineers, October 2013, Volume 2, Number 5: 64-70.

Osmundsen, P., Emhjellen, M., Johnsen, T., Kemp, A. and Riis, C. (2015), "Petroleum taxation contingent on counter-factual investment behavior", *Energy Journal* 36: 1-20.

Osmundsen, P., Løvås, K. and Emhjellen, M. (2017), "Petroleum Tax Competition Subject to Capital Rationing", Proceedings, CESifo Working Paper Series No 6390; https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2951056

Poterba, J. (2010), "The challenge of tax reform and expanding the tax base", *The Economic and Social Review*, 41: 133–148.

Smith, J. L. (2013), "Issues in extractive resource taxation: review of research methods and models", *Resources Policy*, 38: 320–331.

Smith, J. L. (2014), "A parsimonious model of tax avoidance and distortions in petroleum exploration and development", *Energy Economics*, 43: 140–157.

Sylte, J.E., Hallenbeck, L.D., and Thomas, L.K. (1988), "Ekofisk Formation Pilot Waterflood." *SPE-18276-MS*.

Appendix: A model for hydrocarbon production by water injection

To the left in Figure 3 a sketch of an oil reservoir is shown. Water is injected to push oil towards the producer. Ideally, one could imagine a process where water was injected into the aquifer, and the aquifer would rise like a piston to displace the oil. In reality there are complicated flow patterns from the injector to the producer, due to physical heterogeneities (like high permeable zones, faults that could be open or closed, and fractures) and due to a non-uniform fluid pressure distribution.

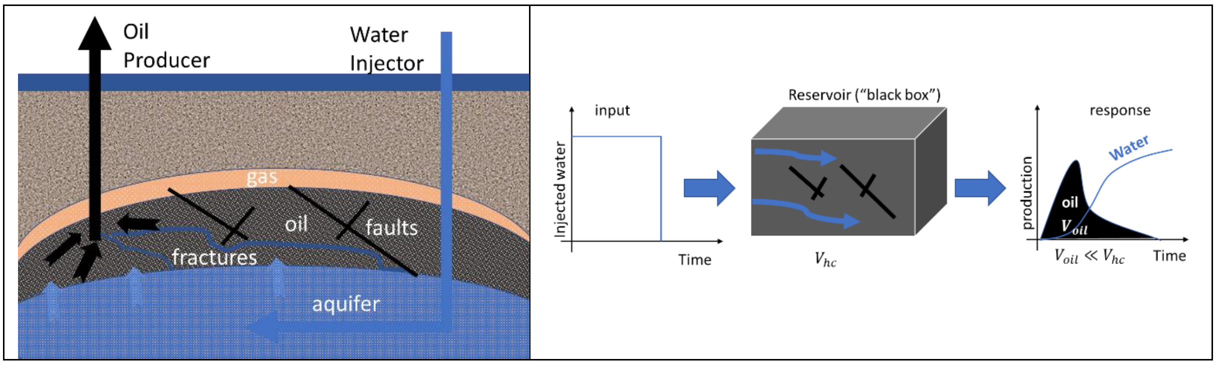

Figure 3. Left: a sketch of an oil reservoir with a gas cap in contact with an aquifer. Water is injected into the aquifer and oil is pushed to the producer. Heterogeneities like fractures, faults or high permeable zones might lead to early water breakthrough in the producer. Right: a simplified model of a reservoir. A pulse of water is injected and a response is measured to the right. Due to suboptimal flow, the water does not contact the full reservoir and only recovers a fraction of the oil.

The subsurface flow pattern is further complicated by the fact that there can be additional pressure support from the underlying aquifer, a gas cap or even reservoir compaction and seabed subsidence. In order to put wells at the optimal locations, one need to know the spatial distribution of oil, water and gas. Therefore, oil companies put a lot of effort into developing reservoir models with a high resolution (millions of simulation cells) to capture the spatial distribution of gas, oil and water. These models might take weeks to run in order to forecast the oil production, and they are constantly being updated in order to be consistent with new information and data. The reservoir models are crucial for making good decisions about field development, such as the placement of infill wells, to gain an optimal production strategy.

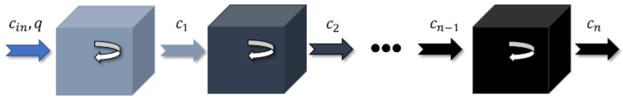

Figure 4: A simple model for a section of a reservoir. All the blocks have equal volumes. Water is injected to the left, and oil and water is produced to the right.

A simple well model

For our purpose, we do not need a detailed description about the spatial distribution of oil and water, but a model that capture a representative field's production over time. We assume that the wells are placed on average in an equally good (or bad) manner. An oil reservoir can then be visualised as a big box, as indicated to the right in Figure 4. Water is injected at one side and a mixture of oil and water is produced on the other side. An injected fluid element may spend a long or a short time in the reservoir before it is produced. If all the injected fluid elements spend the same amount of time in the reservoir, they recover only a small fraction of oil. This is because then most likely, the fluid elements follow the same path, which could happen due to fractures or high permeable zones connecting the injector and the producer. If, on the other hand, there is a wide distribution of residence times of each fluid element, more oil will be recovered, simply because then the water would follow different paths from the injector to the producer and contact a larger volume of the reservoir. The model we propose is based on ideas from chemical engineering, where residence time analyses is used to model chemical reactors, first introduced by Danckwerts in 1952.

We suggest viewing the amount of water in the reservoir and in the produced fluids as a *concentration*. This simple idea greatly reduces the complexity of the problem and leads to a model that to our knowledge has not previously been suggested. If the concentration out of the reservoir is zero, it means that only oil is produced. If only water is injected into the reservoir the concentration into the reservoir is one. By using the residence time analysis formalism, we can then write for any kind of system the concentration of water produced out of a well as a convolution integral (see e.g. Levenspiel 1972):

$$\mathrm{c}_{\mathrm{w}}^{out}(t) = \frac{V_w^{out}(t)}{V_w^{out}(t) + V_{hc}^{out}(t)} = \int_0^t c_w^{in}(t - t')E(t')dt', \tag{4}$$

where V_w^{out}, V_{hc}^{out} is the volume of water and hydrocarbon (oil and gas) being produced out of the well, $c_w^{in}(t - t')$ is the amount of water injected from start to time t, and $q(t)$ is the injection rate of water. The concentration of hydrocarbons produced is then simply:

$$\mathrm{c}_{\mathrm{hc}}^{out}(t) = 1 - \mathrm{c}_{\mathrm{w}}^{out}(t). \tag{5}$$

The function $E(t)$ in equation (4) is called the residence time distribution (RTD), and describes the fraction of injected water that has been in the system between t and $t + dt$. Thus, all the information about the flow paths, like e.g. the amount of fracture vs matrix flow or bypassed hydrocarbons, will be contained in the RTD. The RTD can in principle be estimated from production data. In the following we will use a model to estimate the RTD. Mathematically it is easiest to determine the RTD using a δ-pulse injection. In this context it would mean to inject water for a short period of time, and then switch to oil and measure the fraction of water produced. No operator would do this of course, but we use it as a mathematical trick to estimate the RTD. Then we use equation *(4)* to estimate the concentration of water during a more sensible injection scheme, like constant water injection. The RTD during a pulse injection is given by the following formula:

$$E(t) = \frac{q(t)c_w^{out}(t)}{\int_0^\infty q(t)c_w^{out}(t)}. \tag{6}$$

In the following we will assume that the injection rate in the well is constant. This constraint can be relaxed, but it would not be of any use to us, because we do not know the injection rate of an individual well. Therefore, it makes more sense to assume an average injection rate. In order to proceed we need a model for the reservoir in order to find an expression for the RTD. In the following we will consider a well-mixed tank or a series of well mixed tanks as illustrated to the right in Figure 3 as a model for a producing well. A well-mixed model distributes the fluid elements evenly throughout the tank. A natural question to ask is: why should this simple model be a good model for a producing well? We would argue that the most important physical property to capture of a reservoir is the large-scale heterogeneities. If we consider only one tank between the injector and the producer, it means that there are a great deal of fractures connecting the injectors to the producers as some of the injected water will be produced immediately. By putting many tanks in a series, it will take a longer time to produce the injected water, meaning that much of the oil is displaced before the water is produced. In the limit of infinite number of tanks the production profile will be like a piston displacement, which mimics the behaviour of a homogeneous reservoir. The estimation of RTD for tanks in series can be found in most textbooks (e.g. Levenspiel 1972).:

$$E(t) = \frac{1}{\tau}\left(\frac{t}{\tau}\right)^{n-1}\frac{n^n}{(n-1)!}e^{-\frac{nt}{\tau}}, \tag{7}$$

where n is the number of tanks in series, and τ is a time constant which is related to the *total* volume of all the mixing tanks. The time constant is given by

$$\tau = \frac{V_{hc}}{q}. \tag{8}$$

The physical interpretation of τ is that it is the average well lifetime. The concentration of water is then, according to equation *(4)*:

$$\mathrm{c}_\mathrm{w}^{out}(t) = \int_0^t E(t')dt' = \Gamma\left(n, \frac{nt}{\tau}\right), \tag{9}$$

where $\Gamma(\cdot)$ is the incomplete gamma function. The production rate of hydrocarbons can then be found from equation *(5)*:

$$q_{hc}^{out}(n,\tau,t) = q\cdot\left(1 - c_w^{out}(t)\right) = q\cdot\left(1-\Gamma\left(n,\frac{nt}{\tau}\right)\right). \tag{10}$$

Before we proceed, we will just point out that for $n = 1$, one minus the gamma function simplifies to an exponential function and the production rate from the well is:

$$\mathrm{q}_\mathrm{hc}^{out}(1,\tau,\mathrm{t}) = \mathrm{q}\cdot \mathrm{c}_{hc}^{out}(t) = qe^{-\frac{t}{\tau}}. \tag{11}$$

Note that the mathematical form is equal to an exponential decline model. There are two important differences: (i) this is the production rate for *a single well* (and not the total field), and (ii) the parameters q and τ have a clear physical interpretation. In particular, the product $q\cdot\tau = V_{hc}$, and we expect that the product of the fitted model parameters q and τ should have a numerical value close to the drainage volume of the particular well.

A simple reservoir model to predict field production history

During the lifetime of a producing oil field, oil wells are drilled at different intervals. At the NPD website data are available that describe when a well was drilled for a specific field and which type of well it was (e.g. producer or water injector). In the following we will assume that the hydrocarbon rate from a single well is described by equation (10). We will assume:

- A steady state production and injection rate of fluids throughout the lifetime of the field
- All the wells are connected to the same reservoir
- All the wells drain from separate volumes

From the assumptions above it follows that when the first well is drilled it produces according to equation (8) with the time constant $\tau = V_{hc}/q$, when a new well is put online all the wells produce according to the time constant $\tau/2 = V_{hc}/2q$ etc. Thus, the production from the field at any time is:

$$Q_{\mathrm{hc}}^{field}(\mathrm{t}) = \sum_{i=1}^{N(t)} q_{hc,i}^{out}\left(n, \frac{\tau}{N(t)}, t - T_i\right) H(t - T_i), \tag{12}$$

where $H(t - T_i)$ is the Heaviside step function, and $N(t)$ is the number of wells at time t. Note that for $n = 1$, equation *(12)* is simply a sum of exponential functions, where the decay rate is reduced proportional to the number of wells drilled. The field decay rate $\frac{\tau}{N(t)}$ is then the average field life time. We have fitted the model to the field production data (NPD 2019) for 41 of the largest field on the NCS using the curve_fit() function in Python. The fitted values are plotted against the actual reservoir estimate (NPD 2019) of the reserves to left in Figure 5, and to the right we have tried to make an estimate of the quality of the fit by comparing the area under the production curves (the model and the actual data). The main conclusions that can be drawn are that:

- The model is able for most of the field to fit the total production within ~10%, independent of the size of the field and number of wells drilled
- The estimate parameters in the model matches well the physical parameters

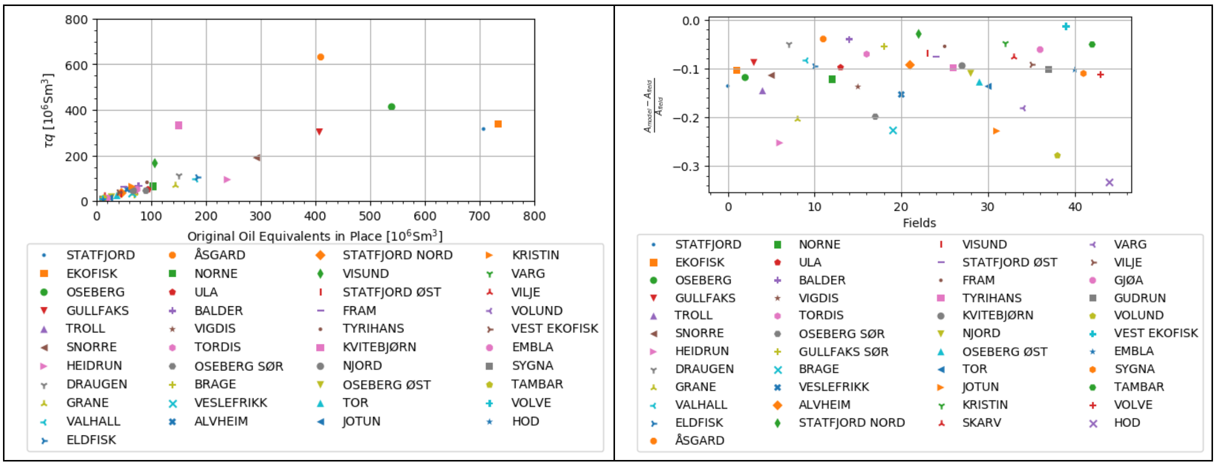

Figure 5: Left: estimated volumes in the ground from the NPD website and the product of the estimated parameters τq (for the n = 1 cases). The Troll field is left out, because the oil equivalents are so large that it is impractical to fit in the same plot. Right: a measure of the quality of the fit. Estimated by calculating the area under the model prediction and the field data, normalized by the area under the field data.

Note that our model so far has neglected effects of pressure decay in the reservoir. We partly take into account the natural energy in the reservoir in the form of pressure support from an underlying aquifer or gas cap, but we neglect the expansion of dissolved gas in oil and the gas cap. However, it is possible to account for these effects by modifying equation *(12)*:

$$Q_{\mathrm{hc}}^{field}(\mathrm{t}) = \mathrm{q} \cdot \mathrm{e}^{-\alpha\mathrm{t}} \sum_{i=1}^{N(t)} c_{hc,i}^{out}\left(n, \frac{\tau}{N(t)}, t - T_i\right) H(t - T_i). \tag{13}$$

In the limit of no water injection, or more correctly no pressure support from an aquifer or gas cap, $\tau \rightarrow \infty$, and equation *(13)* reduces to:

$$Q_{\mathrm{hc}}^{field}(\mathrm{t}) = \mathrm{q} \cdot \mathrm{e}^{-\alpha \mathrm{t}} \tag{14}$$