Commodities Data Analysis

Stephane Ellis

January 2, 2025

Contents

1 Analysis of Oil and Gas Fields on the Norwegian Continental Shelf

1.1 Overview

This section explores oil and gas data for fields located on the Norwegian Continental Shelf. The analysis investigates trends in recoverable oil, investments, and production start dates, aiming to uncover patterns in costs, field sizes, and geographical influences on oil field development.

1.2 Data Overview and Cleaning

The dataset `oil_fields` contains information on recoverable oil, remaining oil, investments, and field characteristics. Initial cleaning steps included:

- Converting production start dates (`producing_from`) to `datetime` format.
- Dropping rows with missing production start dates or latitude data.
- Removing fields with zero recoverable oil to avoid division by zero in investment calculations.

1.3 Key Variables and Visualisations

1.3.1 Recoverable Oil

The total recoverable oil in each field is analysed:

- A scatter plot of recoverable oil against production start dates reveals that older fields tend to have higher recoverable reserves.
- A histogram shows a "fat tail" distribution, with most fields being small but a few outliers, such as *Statfjord* and *Ekofisk*, dominating in size.

1.3.2 Investments and Costs

Key cost metrics were calculated:

- **Investment per extracted oil volume:** Total investment divided by extracted oil (`invest_per_millsm3`).
- **Investment per recoverable oil volume:** Total investment divided by recoverable oil (`invest_per_rec`).

Trends in investments include:

- Larger fields exhibit economies of scale, with lower costs per cubic metre of oil extracted.

- Newer fields are significantly more expensive to develop, as shown by a positive slope in a semi-log scatter plot of investment per extracted volume versus production start dates.

1.3.3 Large Oil Fields

Fields with over 50 million SM^3 of recoverable oil were analysed separately. Key findings include:

- Large fields dominate the total recoverable reserves, with *Statfjord* being the largest.
- Geographical visualisation using latitude and longitude confirms the concentration of large fields in key areas.

1.4 Tax Regime and Investment Incentives

The Norwegian tax regime significantly influences investment in oil fields:

- A marginal taxation rate of 78% (22% corporation tax + 56% resource tax) applies to oil and gas profits.
- Companies can immediately deduct investment costs, enhancing cash flow during the capital-intensive start-up phase.
- These policies expose the government to financial risks if projects fail or oil prices decline.

The tax neutrality assumption could break down under high regulatory risks or prolonged unprofitability, potentially leading to resource misallocation.

1.5 Geographical Analysis

The hypothesis that northern fields are more expensive to exploit was tested:

- A scatter plot of latitude against investment per recoverable oil showed no clear pattern.
- Correlation analysis confirmed a weak positive relationship ($r = 0.02$).
- Linear regression analysis found no statistically significant impact of latitude on costs.

Adding production start dates and recoverable oil as predictors improved the model ($R^2 = 0.139$), showing that later start dates significantly increase costs.

1.6 Insights and Conclusions

The analysis revealed several critical insights:

- Economies of scale in large fields reduce costs per unit of oil extracted.
- Investments in newer fields are more expensive due to technological, regulatory, and geographical factors.
- The Norwegian tax regime plays a pivotal role in incentivising oil field development but introduces significant financial risks to the government.

Future research could explore:

- Depth and number of wells as additional cost factors.
- Temporal changes in cost efficiencies and their drivers.
- Comparative analyses with oil fields in other regions.

2 Accounting Data for Petroleum Firms

2.1 Overview

This section analyses a dataset of Norwegian oil and gas firms, focusing on financial trends and sectoral performance during a turbulent period for the industry. Following the shale revolution (commonly referred to as "fracking"), oil prices fell sharply after 2012, leading to significant restructuring in the Norwegian oil and gas sector. The analysis aims to identify lessons for the future of oil and gas, explore potential shocks to the industry, and evaluate the effects of government policy.

2.2 Dataset and Cleaning

The dataset, `adf`, contains financial and operational information for Norwegian petroleum firms over a seven-year period (2009–2015). Initial cleaning steps included renaming columns to English equivalents, converting numeric industry codes to string format, and splitting the `NACE_code` column into main and sub-sector identifiers. Missing values and redundant columns were addressed as follows:

- The year 2016 was removed due to a complete lack of data.
- Columns with excessive missing values were retained only if they contributed unique, context-specific insights.

2.3 Sector-Level Analysis

2.3.1 Aggregated Sector Data

The dataset was grouped by the main `NACE1` sectors and aggregated by year. Key financial metrics, including total income, operating results, costs, and wage expenses, were calculated. The 2015 data was further analysed, revealing that:

- Sector 70 (*Head offices and management consultancy*) generated the highest total income.
- Sector 6 (*Extraction of crude petroleum and natural gas*) was the second largest, followed by sector 46 (*Wholesale trade excluding motor vehicles*).

2.3.2 Profitability Trends

An analysis of the most profitable sectors over time showed:

- Sector 64 (*Financial services*) and sector 70 consistently outperformed others in terms of profitability.
- Sector 6, representing oil and gas, peaked in profitability in 2011 but declined sharply thereafter.

The top four sectors by profitability were visualised, highlighting significant differences in performance dynamics.

2.4 Firm-Level Analysis

2.4.1 Firm Profitability Rankings

Individual companies were ranked by operating results within their respective sectors. In 2015, the top-performing firms across all sectors were: **DNB Bank ASA**, **Telenor ASA** and **Statoil ASA**.

2.4.2 Profit Margins and Firm Size

Profit margins, calculated as the ratio of operating results to revenue, were analysed:

- Large firms (>3 billion NOK in assets) generally exhibited positive margins, while smaller firms often struggled with negative margins.
- Outliers included **Bayerngas Norge AS** (lowest margin) and **Statoil ASA** (largest assets).

Scatter plots and regression lines illustrated relationships between profit margins, total assets, and other financial metrics.

2.5 Temporal Trends and Drivers

2.5.1 Revenue, Profit, and Liquidity Over Time

Key trends from 2010 to 2015 included:

- Declining total profits and revenues, particularly after 2013.
- Increased reliance on debt financing as equity levels stabilised.
- Cash reserves decreased post-2013, likely due to operational funding needs and debt repayments.

2.5.2 Comparison with Brent Oil Prices

The correlation between Brent crude prices and industry profits was examined. Results indicated:

- Total profits were strongly correlated with oil prices, reflecting the industry's dependency on external market conditions.
- The discrepancy between revenue and profit trends from 2013 to 2014 was attributed to increased operating costs and potential inefficiencies.

2.6 Conclusion and Insights

The analysis highlights critical financial and operational dynamics within the Norwegian oil and gas sector:

- The industry's profitability is heavily influenced by external shocks, such as falling oil prices and structural changes.
- Larger firms are more resilient, with stronger financial buffers and higher profit margins.
- The Norwegian government plays a key role in mitigating losses through tax neutrality laws, effectively subsidising non-profitable firms.

Future research could focus on:

- The impact of government incentives on exploration and profitability.
- Long-term trends in sector concentration and diversification.
- Comparative analysis with other petroleum-producing nations.

3 Analysis of Petroleum Exploration and Financial Data on the Norwegian Continental Shelf

3.1 Datasets

Exploration Data: Data from the Norwegian Petroleum Directorate covering exploratory wells drilled on the Norwegian Continental Shelf over the past decade. The dataset categorises wells into two purposes:

- *Wildcat Drilling:* High-risk exploratory drilling to discover hydrocarbons.
- *Appraisal Drilling:* Post-discovery evaluation of field size and composition.

Additional features include entry and completion dates, well content (e.g., condensate, dry, shows), and the year of operation.

Company Financial Data: Firm-level financials, including total assets, profitability, debt, equity, and other metrics. Data is matched to exploration records via company IDs from the Norwegian Petroleum Directorate and the national register.

3.2 Objectives

- **Exploration Trends:** Examine the temporal dynamics of wildcat and appraisal drilling, particularly in relation to oil price fluctuations.
- **Wildcat Well Analysis:** Assess risk profiles based on drilling outcomes (e.g., dry wells vs. productive wells).
- **Financial Impact:** Investigate relationships between company financial metrics and exploratory drilling decisions.
- **Market Shifts:** Analyse changes in exploration activity across major and minor firms in response to market conditions.

3.3 Methodology

Data Preparation:

- Date fields were converted to `datetime` format to enable time-based analysis.
- Company names were matched across datasets using IDs from national and petroleum registries, with inconsistencies addressed by cleaning and aligning formats.

Exploration Trends: The number of wells drilled per year was plotted, focusing on wildcat wells to capture high-risk activity. Changes in activity were linked to oil price dynamics, such as the 2015 price drop (∼50$/bbl).

Financial Correlations: Regression models were employed to assess the influence of firm financials on drilling activity. Predictors included total assets, profitability, and debt levels.

Risk and Outcomes: Bar charts illustrated the distribution of outcomes (e.g., dry wells, condensate, shows) for wildcat drilling, highlighting the inherent risks.

3.4 Findings

Exploration Activity Trends:

- Exploration peaked from 2010–2014, driven by high oil prices (∼100$/bbl).
- Activity declined post-2015, coinciding with a significant drop in oil prices, reflecting reduced capital expenditure (CAPEX).
- Despite falling prices, 39 wildcat wells were drilled in 2015, possibly due to pre-existing project commitments or a strategy to secure reserves while competition was low.

Wildcat Well Outcomes:

- Majority of wildcat wells were dry (47%).
- Productive wells (e.g., condensate, shows) constituted a minority, underscoring the high-risk nature of exploratory drilling.

Financial Drivers of Exploration:

- Firms with higher total assets drilled more wells, as evidenced by statistically significant regression results.
- Profitability showed no significant direct relationship with the number of wells drilled, suggesting exploration decisions may prioritise long-term strategic goals over short-term profitability.
- Larger firms, such as Equinor, dominated exploration activity, while smaller firms increased their share post-2016.

Market Dynamics and Firm Strategies:

- During high-price periods (2010–2014), firms leveraged debt for growth projects, increasing both equity and exploration.
- Post-2015, firms shifted focus to survival strategies, reducing exploration and CAPEX while maintaining operational efficiency.
- Smaller firms ramped up exploration post-2016, likely capitalising on opportunities left by larger firms reducing activity.

3.5 Conclusion

The analysis highlights the interplay between market conditions, financial strategies, and exploration decisions. While large firms led drilling activity during periods of high oil prices, smaller firms emerged as key players in the low-price environment post-2016. The high-risk nature of wildcat drilling is evident, with nearly half of all wells yielding no significant discoveries. Financial strength, particularly total assets, is a critical determinant of exploration activity, underscoring the importance of long-term resource allocation strategies in the petroleum sector.

4 Analysis of Nordic Electricity Markets and Wind Power Dynamics

4.1 Datasets

Nord Pool Hourly Power Data: Hourly electricity prices (EUR/MWh) for the Danish zones DK1 and DK2, obtained from Nord Pool. The data also includes wind power production (MWh) for the same regions, enabling granular analysis of price and production dynamics.

Net Exchange Data: Hourly net exchange series for Danish DK1 (Western Denmark) and DK2 (Eastern Denmark) regions, representing electricity imports and exports. This dataset facilitates the study of Denmark's trade patterns influenced by wind power production.

4.2 Objectives

- **Stationarity Analysis:** Determine whether electricity prices, wind power production, and net exchange series are stationary or require transformation for time series modelling.
- **Impact of Wind Power on Prices:** Quantify how wind power production, both contemporaneous and lagged, affects electricity prices in DK1.
- **Trade Dynamics:** Analyse the relationship between wind power production and net electricity exports for DK1 and DK2.
- **Addressing Serial Correlation and Seasonality:** Incorporate adjustments to account for autocorrelation and seasonal patterns in residuals.

4.3 Key Methodologies and Concepts

Stationarity and Persistence: The Augmented Dickey-Fuller (ADF) test was used to evaluate stationarity. Non-stationary series, such as DK1 electricity prices, were transformed via differencing to achieve stationarity. Wind power production data was stationary without transformation.

Temporary and Permanent Shocks: Temporary shocks, such as sudden increases in wind production, were modelled for their transient impact on prices, while permanent shocks, such as regulatory changes, illustrated long-term trends.

Serial Correlation: Residuals from regression models were analysed for autocorrelation using ACF and PACF plots. Autoregressive terms (e.g., AR(1), AR(2)) were added to address serial correlation.

Log Transformation and Trends: Log transformations were used to linearise relationships and manage varying scales between electricity prices and wind power. Trends, such as upward drifts in price series, were modelled explicitly when necessary.

4.4 Findings

Impact of Wind Power on Prices:

- A 1 MWh increase in wind power production decreases DK1 electricity prices by approximately 0.34 EUR/MWh.
- The relationship is contemporaneous, with no significant lagged effects observed, indicating an immediate supply-driven price response.

Wind Power and Net Exchange Dynamics:

- Each additional 1 MWh of wind power production in DK1 results in approximately 0.85 MWh of increased net exports.
- Similar to prices, lagged effects of wind power on net exchange dynamics were insignificant.

Seasonality and Residual Diagnostics:

- Serial correlation and seasonality were evident in initial residuals, necessitating the inclusion of autoregressive terms and seasonal adjustments.
- ACF and PACF plots informed the addition of two lags to model serial correlation better.

Persistence and Transformations:

- DK1 electricity prices exhibited high persistence (lag coefficient ≈ 0.95), indicative of a near-random walk. Differencing successfully transformed the series into a stationary form.
- The strong negative relationship between wind power and prices aligns with economic fundamentals: increased supply reduces prices, and wind power's near-zero marginal costs ensure it remains in the market regardless of price levels.

4.5 Conclusion

The analysis underscores wind power's immediate and significant role in shaping electricity prices and trade dynamics in Denmark. Proper handling of stationarity, serial correlation, and seasonality enhances model accuracy and reliability. Wind power's cost structure drives its consistent impact, while its lack of lagged effects simplifies predictive models for market participants.

5 Analysis of Power Market Dynamics Using ARIMA and GARCH Models

5.1 Datasets

ETS Carbon Prices: This dataset includes carbon price data from the European Emissions Trading System (ETS), sourced from the UK think tank EMBER and traded on the EEX exchange. It utilises a "Cap and Trade" mechanism to price carbon emissions, incentivising reductions. Weekly carbon price data was aggregated into monthly averages to align with the analysis of broader energy market trends and long-term seasonal dynamics.

NordPool Power Data: This dataset provides hourly electricity prices (EUR/MWh) for DK1 and DK2. Prices were converted to EUR/kWh by dividing by 1,000 to ensure consistency with ETS carbon price data. The analysis focuses on daily averages, enabling the identification of key trends, seasonal patterns, and their relationships with carbon price fluctuations and renewable energy outputs.

Norwegian Daily Consumption Data: This dataset captures Norway's daily electricity consumption for 2019. It highlights strong weekly and yearly seasonal patterns, particularly reflecting industrial demand variations between weekdays and weekends. The data was analysed to identify mean consumption trends, volatility clustering, and the implications of demand fluctuations on energy market stability.

5.2 Objectives

- **Mean Modelling:** Construct time series models (e.g., ARIMA) for the daily average power prices in DK1 and DK2, assuming constant variance (homoskedasticity). These models aim to forecast expected price levels while focusing on trends, seasonality, and autoregressive patterns in the mean of electricity prices.
- **Volatility Modelling:** Examine the conditional variance of weekly log changes of Norwegian daily consumption to capture heteroskedasticity, particularly volatility clustering. This involves using models such as GARCH to quantify and forecast risk in power markets.

5.3 ARIMA Models for Mean Forecasting

Stationarity of Power Prices: Historically, power prices exhibited $I(1)$-type behavior, with shocks persisting due to reliance on fossil fuels such as coal and oil. In contrast, renewable-driven markets, such as Nordic markets, demonstrate $I(0)$-type behavior, with shocks that are mean-reverting. Stationarity tests on Danish power prices confirmed these renewable-driven markets' tendency towards $I(0)$ properties.

Model Selection: Two ARIMA models were compared:

- **ARIMA(0,1,1)** ($I(1)$): Flat forecasts with persistent shocks that reflect non-stationary dynamics.

- **ARIMA(2,0,0)** ($I(0)$): Mean-reverting forecasts.

The ARIMA(2,0,0) model was chosen due to its logical consistency with renewable power markets and better AIC/BIC values.

5.4 ARMA-GARCH and SARIMA-GARCH Models for Volatility Analysis

Seasonality in Daily Consumption: Weekly and yearly seasonal patterns were identified in daily electricity consumption data, with higher average consumption on weekdays compared to weekends. These patterns necessitated incorporating weekly seasonality when analyzing demand-driven volatility. Seasonal ARIMA (SARIMA) models were utilized to capture these periodic fluctuations and improve the accuracy of time series models.

ARMA-GARCH Findings: An ARMA(1,1) model was fitted to weekly log changes in daily electricity consumption, with GARCH(1,1) applied to the residuals to capture volatility clustering. The GARCH component effectively modeled conditional heteroskedasticity, as evidenced by a significant α coefficient, indicating that recent shocks (e.g., large changes) strongly influenced current volatility. The persistence parameter (β) was close to 1, suggesting that volatility decayed gradually after.

Forecasts produced by ARMA-GARCH revealed periods of elevated volatility following market shocks. However, significant residual autocorrelations at lags that were multiples of 7 revealed unmodelled weekly seasonality. This underscored the necessity of incorporating seasonality into the model to improve the accuracy of forecasts in datasets with periodicity.

SARIMA-GARCH Findings: To address the deficiencies of the ARMA-GARCH model, a SARIMA(1,0,1)(0,0,1)[7] model was fitted to the log changes, with GARCH(1,1) applied to its residuals. This approach integrated weekly seasonality directly into the ARMA component. The SARIMA-GARCH model captured volatility clustering while resolving the residual autocorrelation issues observed in the ARMA-GARCH framework. The persistence parameter (β) for conditional volatility was moderate and statistically insignificant, suggesting a faster decay of volatility compared to ARMA-GARCH. Additionally, the model revealed seasonal peaks in conditional variance, particularly during winter and spring, reflecting higher uncertainty in electricity consumption during these periods.

5.5 Comparison of Models

ARMA-GARCH: The ARMA-GARCH model effectively captured short-term volatility dynamics, with significant clustering of conditional volatility in residuals. However, its inability to model seasonality led to biased forecasts and residual autocorrelation.

SARIMA-GARCH: The SARIMA-GARCH model addressed the limitations of ARMA-GARCH by incorporating weekly seasonality, leading to improved residual diagnostics and more accurate representation of volatility dynamics. Its ability to account for cyclical patterns made it particularly suitable for markets influenced by periodic fluctuations, such as daily electricity consumption.

Forecast Analysis: Forecasts from ARMA-GARCH exhibited wide confidence intervals and missed recurring seasonal patterns, while SARIMA-GARCH produced tighter intervals and better captured the cyclical nature of volatility. The improved representation of conditional variance under SARIMA-GARCH highlights the importance of seasonality in modeling renewable-driven electricity markets.

5.6 Conclusion

The analysis demonstrated the importance of incorporating stationarity, seasonality, and volatility clustering in modelling power market data. While ARIMA models captured mean trends effectively, GARCH models were essential for understanding volatility dynamics. SARIMA-GARCH models provided the most comprehensive representation by addressing both seasonality and conditional variance, aligning with the characteristics of renewable-driven markets.