St. Petersburg University
Graduate School of Management

[Master in Corporate Finance]

IMPLEMENTATION OF SIMULTANEOUS EQUATIONS MODEL TO FORECASTING RESIDUAL INCOME OF RUSSIAN OIL COMPANIES

Master's Thesis by the 2nd year student
Sennikov Igor

Research advisor:
Associate Professor, Garanina T. A.

St. Petersburg

[2016]

Statement about the independent character of the master thesis

I, Sennikov Igor, (second) year master student, Master in Corporate Finance, program «Management», state that my master thesis on the topic «Implementation of simultaneous equations model to forecasting residual income of Russian oil companies», which is presented to the Master Office to be submitted to the Official Defense Committee for the public defense, does not contain any elements of plagiarism.

All direct borrowings from printed and electronic sources, as well as from master theses, PhD and doctorate theses which were defended earlier, have appropriate references.

I am aware that according to paragraph 9.7.1. of Guidelines for instruction in major curriculum programs of higher and secondary professional education at St.Petersburg University «A master thesis must be completed by each of the degree candidates individually under the supervision of his or her advisor», and according to paragraph 51 of Charter of the Federal State Institution of Higher Education Saint-Petersburg State University «a student can be expelled from St.Petersburg University for submitting of the course or graduation qualification work developed by other person (persons)».

(Student's signature)
26/05/2016 (Date)

Заявление о самостоятельном характере выполнения выпускной квалификационной работы

Я, Сенников Игорь, студент второго курса магистратуры направления «Менеджмент», заявляю, что в моей магистерской диссертации на тему «Применение системы одновременных уравнений для прогнозирования остаточного дохода российский нефтяных компаний», представленной в службу обеспечения программ магистратуры для последующей передачи в государственную аттестационную комиссию для публичной защиты, не содержится элементов плагиата.

Все прямые заимствования из печатных и электронных источников, а также из защищенных ранее выпускных квалификационных работ, кандидатских и докторских диссертаций имеют соответствующие ссылки.

Мне известно содержание п. 9.7.1 Правил обучения по основным образовательным программам высшего и среднего профессионального образования в СПбГУ о том, что «ВКР выполняется индивидуально каждым студентом под руководством назначенного ему научного руководителя», и п. 51 Устава федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет» о том, что «студент подлежит отчислению из Санкт-

Петербургского университета за представление курсовой или выпускной квалификационной работы, выполненной другим лицом (лицами)».

(Подпись студента)
26/05/2016 (Дата)

Abstract

Master Student's Name	Igor Sennikov
Master Thesis Title	Implementation of simultaneous equations model to forecasting residual income of Russian oil companies
Faculty	Graduate School of Management
Main field of study	Corporate Finance
Year	2016
Academic Advisor's Name	Garanina T. A.
Description of the goal, tasks and main results	Main goal is development of simultaneous equations model to forecasting residual operating income and verification on the sample of Russian oil companies. Main tasks - conduct literature review on accounting based valuation and choose the applicable method of valuation; extend the chosen valuation model, taking into account the features of oil industry in Russia; collect the data for model estimation; applicate the Simultaneous Equations Model to the extended valuation model; develop Visual Basic Application for forecasting purposes of residual operating income of Russian oil companies; make managerial implications of the obtained results. As a result of current research, there was elaboration of forecasting model and Virtual Basic Application development.
Keywords	Value Based Management, Valuation, Oil, Residual Income, Simultaneous Equations.

Аннотация

Автор	Игорь Сенников
Название магистерской диссертации	Применение системы одновременных уравнений для прогнозирования остаточного дохода российский нефтяных компаний
Факультет	Высшая Школа Менеджмента
Направление подготовки	Корпоративные Финансы
Год	2016
Научный руководитель	Гаранина Т. А.
Описание цели, задач и основных результатов	Основная цель данной работы заключается в разработке модели одновременных уравнений и проверка модели на примере российских нефтяных компаний. Задачи – изучение литературы и выбор подходящей модели оценки; расширение существующей модели оценки, учитывая специфику нефтяной отрасли; сбор данных для выполнения эконометрических расчетов; применение системы одновременных уравнений для расширенной модели; разработка программы в среде VBA; вывод управленческих решений, основанных на применении разработанной модели. В результате проведенного исследования, была разработана улучшенная модель для оценки российских нефтяных компаний, а также разработано приложение в среде VBA для оценки.
Ключевые слова	Ценностно-ориентированный менеджмент, Оценка, Нефть, Остаточный доход, Одновременные Уравнения.

Table of Contents

Statement about the independent character of the master thesis 2
Abstract...4
Introduction.. 7
Chapter 1. Accounting approaches to company’s valuation...9
1.1 Value-Based Management..9
1.2 Residual Income Models..13
1.3 Simultaneous Equations Model... 18
Chapter 2. Methodology..25
2.1 Estimation Model Development and Data Description... 25
2.1.1 Application of Simultaneously Equations Model..25
2.1.2 Data and Descriptive Statistics...29
2.1.3 Application Three-Stage Least Squares Initial Estimation..36
2.1.4 Extended model..37
2.1.5 Generalized Two-Stage Least Squares Estimation...43
2.2 Practical implications..46
2.2.1 Application of Panel Vector Autoregression.. 46
2.2.2 Virtual Basic for Application Program Description... 49
Conclusion..52
References.. 53

Introduction

Companies' valuation is an essential part of investing activity. There is special discipline, which main purposes are to create value and to measure it precisely – value based management. Current masterwork is devoted to work with second part – valuation process.

There are different valuation model with different techniques and assumptions. Also, from industry to industry, there is a necessity to apply different valuation techniques and use industry-specific assumptions. In this masterwork, main valuation model are considered and highlighted their main features. In present time, almost all valuation model are quite universal, which could be good for general valuation purposes. On the other hand, there is the one main disadvantage - they are not always effective in implementation for different industries. Main accent is done on specific, narrow-specialized valuation model for industry valuation purposes.

Next important question, discussed in this masterwork - choosing the suitable and interesting industry, which is open for improvements in valuation sphere due to its specific operations. Russian oil industry was chosen for that role. In current masterwork eight biggest oil companies are considered, still not the whole industry, so it will be more accurate to mention companies' valuation.

There was a separation an oil companies from gas extracting, because of existence of some significant differences in processes and performance indicators. Actually, one important mention about not working with whole industry is that this masterwork is focused only on upstream operation as on the most important part for oil companies. Only upstream oil companies have specific performance indicators like oil production or available oil reserves.

Due to working with production and exploration oil processes, main valuation model, on which whole research is focused on – residual operating income model. This choice was done after deep investigation of existing valuation models.

Next important research findings in the current masterwork is deep investigation and study of simultaneous equation model, as a possible key agent for improving of model accuracy. Based on previous researches done with study of simultaneous equations model, there found a fact that this model could be implemented and used for improving of residual operating income model – specially one unknown variable which need to be forecasted – residual income.

Residual operating income model was broadly studied and widely used for different research purposes. Moreover, simultaneous equations model is quite known tool, especially for economic purposes. But combination of these two model for improving of accuracy purposes was never studied before. Also, if we are talking about implementation and development of different valuation models, they are done predominantly for set of companies from different countries or especially for USA companies. There is the main relevance of current masterwork.

As further improvement, researches done is this work can be extended and use not only for Russian oil companies, but for oil companies based in different countries. Of course, for these purposes, special assumption are needed to be done.

To summarize purposes of current work, this master work is dedicated to study of Russian oil industry. The main purpose is to create a valuation model specifically for this industry. This choice was done not accidently: oil industry has a significant share in formation of Russian budget.

While performing research, very important is to precisely set a research goal, which will be the main result of the work to achieve. In our case it is:

- Development of simultaneous equations model to forecasting residual operating income and verification on the sample of Russian oil companies.

It is very hard to achieve the research goal without setting a plan of steps with clear research objectives. They are the next:

- Conduct literature review on accounting based valuation and choose the applicable method of valuation.
- Extend the chosen valuation model, taking into account the features of oil industry in Russia.
- Collect the data for model estimation.
- Applicate the Simultaneous Equations Model to the extended valuation model
- Develop Visual Basic Application for forecasting purposes of residual operating income of Russian oil companies.
- Make managerial implications of the obtained results.

Results of current research done has a very strong managerial implication for a case of working with real oil companies.

Chapter 1. Accounting approaches to company's valuation

1.1 Value-Based Management

Before investing money into some company and even think about it, we need to understand the main purpose of each company. However, it is not so easy to formulate this aim precisely. Nowadays, there are two main theories.

- *Shareholders theory*. According to Jensen (Jensen, 2001), main purpose of each company is to increase its value for shareholders. In this theory, author means also creditors, and by increasing its welfare, society receives positive impact.
- *Stakeholders theory*. Second important theory is presented by Wallace (Wallace, 2003) and implicates not only shareholders and creditors, but all people somehow involved in company's operations and functioning. This theory present function of the company as multi-factor.

Determining and finding this firm's value is the main purpose of Value-Based Management (VBM). For successful achievement of this aim, we need to clearly understand, what is included is VBM and what steps we need to undertake for goal reaching. Understanding of this mechanism help us in in determining crucial valuation points. Pursuant to Copeland (Copeland, Koller, Murrin, 1995), structure of VBM can be drawn as next:

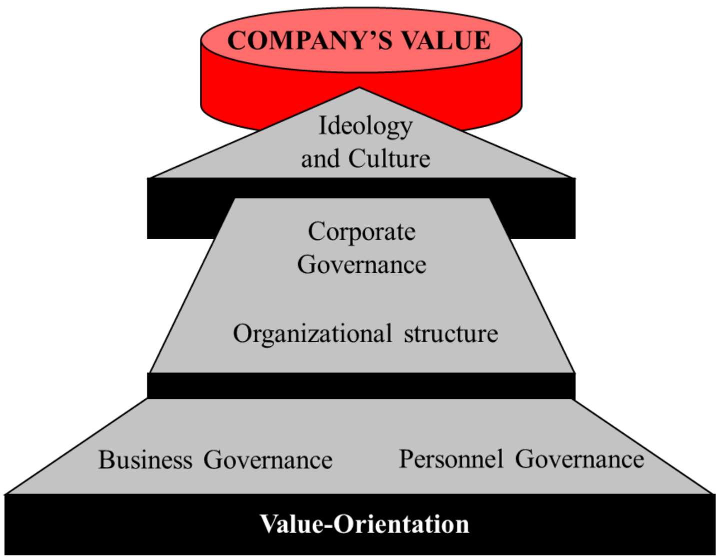


Figure 1.1 Structure of VBM

As we can see from the picture, effective increasing and value determining is not possible without clear and effective management. More clear definition of difference between "governance" and "management" is given by one of professors from Graduate School of Management (Bukhvalov, 2005). We will not stop on governance and management and will go more into valuation part.

In this masterwork, VBM was described as a necessary introduction for valuation-strategy for company. According to Volkov (2006), Value-Based Management consists of four main modules: valuation, strategy, finance and corporate governance.

- *Valuation module*. This module presumes choosing of appropriate valuation model and all necessary assumptions and procedures for determining company's value. In addition, it implements monitoring system and crucial drivers for new value creation.
- *Strategy module*. Determines precise link between company's value and its business-strategies.
- *Finance module*. Describes company's financial policies, which directed to value creating.
- *Corporate Governance module*. Aimed on negotiations of shareholders' and managers' interests.

This masterwork is aimed for valuation purposes, so we will concentrate only on one module. We will examine valuation module in details, because valuation module is a core architecture for VBM. Implementing this module means to consider and make necessary decisions in its three parts (Volkov, 2006):

- Which model to choose?
- Which core indicators to choose?
- How to implement it through value-drivers creation?

This masterwork will be built according to these steps. But before making any decision about model, we need first to determine main purposes of all valuation models. It is quite obvious that we need to determine company's value. First, we need to answer the question: what is the company's value and on which type we need to be more focused? We have three main indicators:

- *Book Value*. This indicator based on difference between total assets and total liabilities – very common measure could be seen in every Balance Sheet. Book Value measure historical value of assets, not trading one. In other words, book value of company is that its shareholder should receive in case of liquidation. However, as assets could be trading at different value (actually, market value), so this case is unlikely. Nevertheless, we could use book value as useful benchmark for determining its value, even having recent study with loud name "Size Matters, Book Value Does Not!" (Lambert, Hubner, 2014), which reconsider whole usage of Fama-French CAPM model (Fama, French, 2004).

- *Market Value.* This is indicator of what investors are willing to pay for company's assets. This indicator could differ significantly for book value and depends more on investor's expectations. Defining the market value of private companies is quite difficult task, because they are not publically traded. Many adjustments need to be done or many peers need to be taken to define this market value precisely. Recent study (Destri, Picone, Mina, 2012) offers implementation of new technologies for determining market value. We will not stop on them, because the only thing we are interested in this masterwork about market value – determining, is our target firm overvalued or undervalued.
- *Fundamental Value.* Also, can be called as "intrinsic value" (Penman, 2010) or "warranted value" (McTaggart, Kontes, Mankins, 1994). It is true actual value of a company, the most interesting kind of value for us. The purpose of many valuation tool is to precisely derive this company's value, which can be quite complicated not only for private companies. Fundamental value, opposite to market, is driven by internal expectations, not public. Deep analysis of main differences is presented in Brainard's fundamental work (Brainard, Shapiro, Shoven, 1990). We will not stop on describing of its features, just highlight that this firm's value is the most interesting for us.

When we are speaking about company's value, we are most interested in fundamental value. Fundamental value is only "true" indicator of company's performance. According to Damodaran (2003), company can have fundamental value if only creates flows to shareholders or stakeholders. Nevertheless, choosing market, book or fundamental value will not cancel the main balance equation. We will write it for fundamental value, as the most interesting indicator for us:

$$V_A(V_F)=V_E+V_D \quad , \tag{1.1}$$

Where:

V_A - fundamental value of assets (or V_F - fundamental value of firm),

V_E - fundamental value of equity,

V_D - fundamental value of debt.

According to the formula (1), we can derive two main approaches to the fundamental valuation of equity.

- *Operational Approach.* Following this method, we determine fundamental value of the firm as sum of discounted flows, associated with firm's assets. Then it is necessary to subtract discounted outflows, associated with debt. By doing this, we receive fundamental value of company.
- *Capital Approach.* Fundamental value of firm determined as discounted future flows to shareholders.

Main difference in these approaches is in underlined assumptions and the process of calculation of book value of debt. Here, in line with operational approach for calculating the value of debt we need to discount all possible outflows, associated with debt. However, if we need to calculate amount of debt using capital approach, it will be sufficient just to take book value of debt.

Next interesting thing is that almost all valuation models can be viewed as equivalent under certain assumptions. Volkov and Rudenko (2009) highlighted core assumptions for an achieving of equivalence between certain models:

- *Discounted free cash flow model (DFCFM) and residual operating income model (ReOIM) equivalence.* Can be achieved by using the same discount rate and general assumptions about calculation the terminal value (for example, by applying constant earnings before interest or taking constant return on net assets).
- *Discounted free cash flow model (DFCFM) and residual earnings model (REM) equivalence*. Can be achieved by implementing two conditions: using general assumption about calculation of terminal value and by calculating of WACC based on fundamental value of financing sources for DFCFM.
- *Residual operating income model (ReOIM) and residual earnings model (REM) equivalence*. Can be achieved by calculating of weighted average cost of capital based on fundamental value of financing sources for ReOIM.

Having two main approaches underlined, now we need to define what kind of flows we may use. Here we have:

- Cash flows
- Residual Income flows

For better understanding and highlighting the core valuation models, table below is displaying the most common approaches:

Table 1.1 Classification of valuation models (Volkov, 2006)

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Approaches to equity valuation</td></tr>
<tr><td>Operational</td><td>Capital</td></tr>
<tr><td rowspan="4">Types of flows</td><td rowspan="2">Cash Flows</td><td colspan="2">Discounted Cash Flows Models (DCFM)</td></tr>
<tr><td>Discounted Free Cash Flows Model (DFCFM)
$V_E^{FCF} = \sum_{j=1}^{\infty} \frac{FCF_j}{(1+k_W)^j} - D_{BV}$
FCF - free cash flow, k_W - weighted average cost of capital,
D_{BV} - book value of debt</td><td>Dividend Discount Model (DDM)
$V_E^{DDM} = \sum_{j=1}^{\infty} \frac{d_j}{(1+k_E)^j}$
d - dividends, k_E - cost of equity</td></tr>
<tr><td rowspan="2">IncomeResidual</td><td colspan="2">Residual Income Model (RIM)</td></tr>
<tr><td>Residual Operating Income Model (ReOIM)
$V_E^{\mathrm{Re}\,OIM} = \left[NA_0 + \sum_{j=1}^{\infty} \frac{ReOI_j}{(1+k_W)^j}\right] - D_0$
NA - net assets, $ReOI$ - residual operating income</td><td>Residual Earnings Model (REM)
$V_E^{REM} = E_0 + \sum_{j=1}^{\infty} \frac{RE_j}{(1+k_E)^j}$
RE - residual earnings, E - equity</td></tr>
</table>

After consideration the most common valuation model, it is time to move further for identification of main differences and features. Next part is dedicated to a closer look at residual income models.

1.2 Residual Income Models

Company valuation is a crucial field for modern corporate finance. The aim of each company is to increase its market value; therefore, there is a need of precise valuation models. Who is a target for “holding” company’s value – stakeholders or shareholders – another field for discussions (Jensen, 2001). This master work is dedicated to valuation models. Plenty of them were invented in recent years. The most accurate of them were described with implementation for Russian market by Rutgayzer (2007).

One important group, widely used by many financiers is valuation methods based only on operating profitability and growth prospective. This paper focuses on one method from this group - Residual Income model (RIM) (or Residual Earnings model), developed by Ohlson (1995). He mentioned and proved relationship between accounting data such earnings, book value, dividends and market value of a company.

Residual Income model was popularized by Penman (2007) in his textbooks dedicated to financial analysis. In Russia, this method was studied by Volkov (2006).

From the very beginning, this valuation method found many researchers, who claimed this method as superior by its accuracy. First, RIM was tested using large simple of firms during 15-years period prior to study (Bernard, 1995). Author concluded this model as effective and more preferable to use for short time period predictions. Many later articles were dedicated to analyzing of this model, implementing it for different companies and different industries. Penman and Sougiannis (1998) support the superiority of accounting-based methods over methods, which use cash flows. Then, Francis at al. (2000) stated better accuracy of RIM, comparing to dividend discount model (DDM) and discounted cash flow model (DCF). Later, Lundholm and O'Keefe (2001) stated that there are should be no difference between DCF and RIM, because they are using the same underlying assumptions.

Nevertheless, RIM is universal method and could not provide same accurate results for different industries. Perek (2012) in his study did not find significant difference between DCF and RIM for Turkish textile industry.

Next step in improving accounting based models was done by Bergmann (2013). Author developed her own model for improving income forecast as an input into valuation models. The main purpose was to use publicly available information, instead of using just information, provided by financial analysts. Barth (2005) used this information as a component to his equity valuation model, specifying it as "other information". He developed three linear information valuation models (LIM) and tested them. His model were based on earnings too, as in Bergmann's (2013) work. The main difference was that Barth (2005) tested influence of disaggregated earnings on equity forecast prediction. The accuracy of each model were tested by comparing errors, without using analysts' forecasts as benchmarks.

However, earlier, Ohlson (2001) in his later work was offering not to avoid analysts' forecasts, but to use them as a proxy for new-derived models. Cheng (2005) supported this point of view, mentioning that some useful information like industry characteristics can be found only in analysts' forecasts. Not all researchers support using analysts' forecasts as benchmarks, or somehow trying to implement them into valuation model. Richardson and Tinaikar (2004) earlier wrote that it "is a tautology that one can outperform a parsimonious relation when analyst forecasted abnormal earnings are added to the model containing only accounting realizations" (p. 228).

To find a compromise, Bergmann (2013), in addition to her simultaneous equations model (SEM) used not only method of comparing errors, but benchmarks as autoregressive model and random walk model. As a result, model of income forecasting, more stable and accurate for crisis years, were developed. In this paper, inspired by work of authors from Graduate School of Management (Bukhvalov et al. 2012), new model derives from universal one and applies for reality of Russian oil industry.

This masterwork is dedicated to one question – how to improve Residual Income Valuation Model (RIM or RIV) and make more precise for current industry. First, we need to determine this model and decide which weak parts could be improved. Significant contribution to creation of this model was done by Ohlson (1995). Nevertheless, first attempts to formulize and derive this method were done much earlier – by Edwards & Bell (1961) and then, later by Peasnell (1982). In the simplest form, we could write this formula as:

$$V_0 = BV_0 + \sum_{t=1}^{\infty} \frac{RI_t}{(1+r)^r} , \qquad (1.2)$$

where:

V_0 - value of a company at time zero,

BV_0 - book value of company at time zero,

RI_t - residual income at time t ,

r - interest rate.

One important note – the formula (1.2) could be applied for mature company only, with a constant growth rate g . Next step is an adjusting the formula for company, having a fact this company is still growing. Here we need to introduce "terminal value" – a company's present value of all cash flows from some point of time in future (time m), when we, as analysts, expect company will have constant growth rate. Different approaches could be done to this calculation, including most popular one – Gordon Growth Model (Gordon and Shapiro, 1956). More complicated and formal approach could be done according to Levin and Olsson (2000). In our masterwork we will not stop on deep analysis of these methods, just rewrite RIV formula in the next way for still growing companies:

$$V_0 = BV_0 + \sum_{t=1}^{m-1} \frac{RI_t}{(1+r)^r} + \frac{T_m}{(1+r)^{m-1}} , \qquad (1.3)$$

where:

V_0 - value of a company at time zero,

BV_0 - book value of company at time zero,

RI_t - residual income at time t ,

r - interest rate,

T_m - terminal value at time m .

Therefore, the main purpose of this model is to determine the "true" value of a company for present moment. Looking at this formula and having assumption like constant interest rate r , one important question arises – how correctly forecast residual income? To answer this question, we need to dig into structure of residual income and understand the algorithm for its deriving.

Here we need to make some distinguishing between residual income and residual earnings. For that purpose, we need to determine components of earnings. According to Penman (2007), three crucial components are in the scheme:

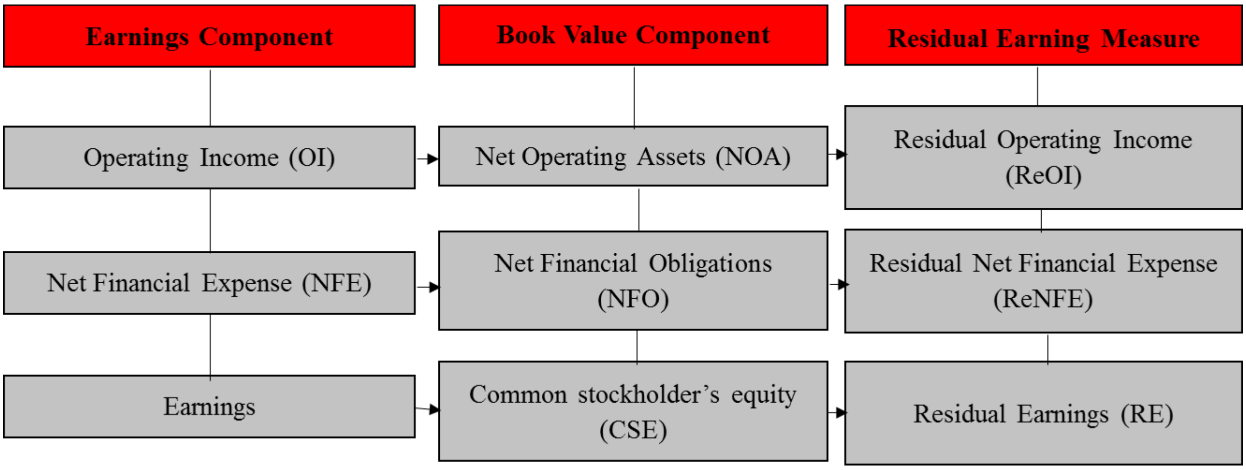


Figure 1.2 Components of Earnings and Book Value

First, let us rewrite residual earning model for equity valuation:

$$V_0^E = CSE_0 + \frac{RE_1}{(1+r_E)^1} + \frac{RE_2}{(1+r_E)^2} + \frac{RE_3}{(1+r_E)^3} + \ldots \quad (1.4)$$

where:

V^E - value of equity,

CSE - common shareholders' equity,

RE - residual earnings,

r_E - cost of equity.

Second, we need to write equation, how we receive residual earnings:

$$RE_t = Earnings_t - r_E \cdot CSE_{t-1} \quad (1.5)$$

Where:

CSE - common shareholders' equity,

RE - residual earnings,

r_E - cost of equity.

REM anchors valuation of equity on the book value of equity. The result is forecasted earnings minus required return on book value. Now, if we take residual earnings from net operating assets, we receive residual operating earnings (ReOI). Here we consider the most important valuation model for current masterwork. The formula is given below:

$$V_E^{\mathrm{Re}\,OIM}=\left[NA_0+\sum_{j=1}^{\infty}\frac{ReOI_j}{(1+k_W)^j}\right]-D_0 \tag{1.6}$$

Where:

NA - net assets,

$ReOI$ - residual operating income,

D - debt,

k_W - weighted average cost of capital.

As for improving this model for oil companies' valuation, we are most interested in residual operating income, which calculated as:

$$ReOI_t=OI_t\cdot(1-tax)-WACC\cdot NOA_{t-1}, \tag{1.7}$$

For further work, we need to determine operating income and derive it with a usage of sales and operational expenditures:

$$ReOI_t=(SALES_t-OPEX_t)\cdot(1-tax)-WACC\cdot NOA_{t-1}, \tag{1.8}$$

Where:

$SALES$ - total revenues,

$OPEX$ - operational expenditures,

$WACC$ - weighted average cost of capital,

NOA - net operating assets.

The formula (1.8) is the most important for us and improvement of its variables is main purpose of this masterwork.

This part was dedicated to model choosing and explaining the main features of the decision. Next part will consider the simultaneous equations model and uncover, why this model is so important for developing of residual operating income.

1.3 Simultaneous Equations Model

What is simultaneous equations model (SEM) and how it is different from structural equations model? SEM is one form of structural equations, also SEM have a "structural" part in their building form. A lot of attention were paid to study of SEM by Cowles Commission [1], the foundation at Yale University, which purpose is conducting various of economic researches. This

1 http://cowles.yale.edu/ Official website of Cowles Commission

foundation is very important for econometric development and its researches are highly cited in many scientific and econometric articles.

The difference between SEM and the whole group of structural equations were underlined by Reiss and Wolak (2007). Simultaneous equation model is the model, where several equations with mutually or jointly dependent endogenous variables. For better understanding of that kind of structural equations, where is the best way to present the most popular economic functions: supply and demand. McFadden (1999) firstly proposed this simple explanation.

Demand function:

$$Q_t^d = \alpha_0 + \alpha_1 P_t + \alpha_2 Y_t + u_{1t};\ \alpha_1 \alpha_2 < 0 \qquad (1.9)$$

Supply function:

$$Q_t^s = \beta_0 + \beta_1 P_t + \beta_2 Z_t + u_{2t};\ \beta_1 \beta_2 > 0 \qquad (1.10)$$

Equilibrium condition:

$$Q_t^d = Q_t^s \qquad (1.11)$$

Where:

Q_t^d - quantity demand,

Q_t^s - quantity supply,

P - price,

Y - disposal income,

Z - raw materials costs,

t - time.

By combining demand function (1.9), supply function (1.10) and the main equilibrium (1.11), we receive a system of simultaneous equations. In received system both supply and demand functions needed to be found simultaneously.

In this masterwork two types of variables will be mentioned several times: endogenous and exogenous. For further work, there is a need of precise definition of them.

- Endogenous variables
  If we are talking about endogenous variables in a system of simultaneous equations model, those variables are jointly depended. Their values are determined within the model. In other words, we need to find exactly their value.
- Exogenous variables
  Those variables in a context of the system of simultaneous equations have specified values from the very beginning. They are predetermined outside the system, in other words, they are given. They often known as non-stochastic.

The figure below is digging into further structure of endogenous and exogenous variables:

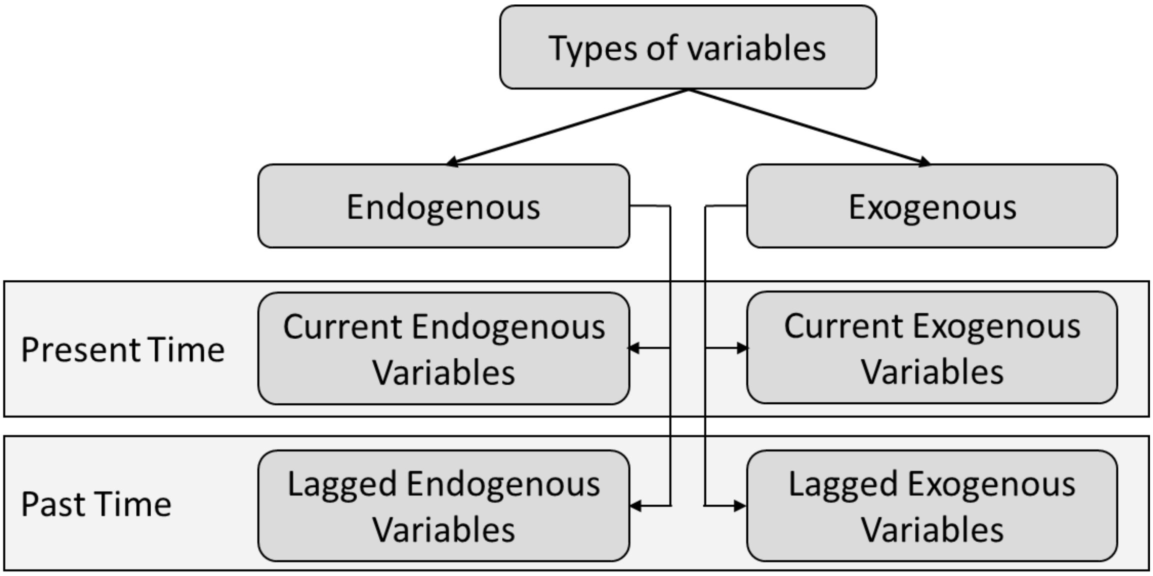


Figure 1.3 Types of Variables in SEM

After defining endogenous and exogenous variables, there is one important addition of SEM is complete or incomplete condition. We need to define endogenous variables and for that purposes we have an amount of equations. Let us consider the next income/consumption system of simultaneous equations:

$$C_t = \beta_0 + \beta_1 y_t + u_t$$
$$Y_t = C_t + I_t \quad (1.12)$$

Where:
C - consumption,
Y - income,
I - investment,
t - time,
u - error term.

Further, there is a necessity of precise bullet pointing the difference of SEM not from structural equation, but from single equations model. Features of structural equation modelling (as group, includes simultaneous equations) first was highlighted by Karimimalayer and Anuar (2012). Earlier, a guide for structural equations modelling was compiled by Schumacker and Lomax (1996). The table below shows the main differences:

Table 1.2 Main difference of SEM from single equation model.

Single Equation Model	Simultaneous Equations Model
In this type of model, single relationship	Relationship inside this type of model

between variables is presented.	includes various interdependencies between variables.
One equation.	Two and more equation without any limitations in number.
Ordinary least squares (OLS) estimation may be used.	No possibility of using OLS, because of interdependencies presence.
Parameters for ingle equation can be estimated just by a one variable	Using of just one equation in a system and usage of just one variable will not help us in overall system estimation.

Equations in SEM can be presented in different forms:

- *Structural form.*
  This form show the complete system of equations. Here, the structure of all relationships between variables is fully presented. Endogenous variables in this case are presented as functions of other endogenous variables, exogenous variables and disturbance terms.
- *Reduced form.*
  The other type of expressing SEM, a model, where endogenous variables are expressed as explicit function. This function is taken of predetermined variables. In other words, here endogenous variables not expressed through other endogenous variables, but only with predetermined exogenous variables and stochastic disturbance.

When we are talking about simultaneous equations and forms of their expression, there is a need of implementing and explanation of recursive model.

- *Recursive model.*
  A way of expression the equations with a special order of equations. Here, the first equation includes only those variables, which are predetermined, in a right side. Next, second equation includes not only its predetermined variables in the right side, but also endogenous variables from first equation. All further equations expression follows the same logic and depends on amount of equations in the system.

As formula, recursive form of SEM could be presented as next:

$$\begin{aligned} Y_1 &= f\left(x_1, x_2, \ldots, x_k u_1\right) \\ Y_2 &= f\left(x_1, x_2, \ldots, x_k Y_1, u_2\right) \\ Y_3 &= f\left(x_1, x_2, \ldots, x_k Y_1, Y_2, u_3\right) \\ Y_4 &= \ldots\ldots \end{aligned} \tag{1.13}$$

Next, after consider all possible types of SEM and different ways of expression, it is time to move to visual and understandable way of showing the model – graphical. But first, we need to define the generally accepted notation of SEM (Bollen, 1989):

Table 1.3. Notation for Simultaneous Equation Models (Bollen, 1989)

Vector/Matrix	Definition	Dimensions
Variables		
x	Endogenous variables	$p \times 1$
y	Exogenous variables	$q \times 1$
ζ	Disturbance terms/errors	$p \times 1$
Coefficients		
Γ	Coefficient matrix for the exogenous variables; effect of exogenous variables on endogenous variables; direct effects of x on y	$p \times q$
B	Coefficient matrix for the endogenous variables; effect of endogenous variables on endogenous variables; direct effects of y on y	$p \times p$
Covariance Matrices		
Φ	Covariance matrix of exogenous variables, x	$q \times q$
Ψ	Covariance matrix of exogenous variables, ζ	$p \times p$

Next important thing about SEM is the pictorial way of representation. It called path diagrams. Having theoretical model, for better understanding it is important to see, how variables are related to each other. Usage of path diagrams follows some basic rules:

- Observed variables are shown in rectangles
- Direction of influence are denoted by usage of single-headed arrows
- Double-headed arrows depict covariance, which is not uncovered in the model
- Unobserved and latent variables usually depicted unenclosed, but also could be shown in ovals.

Let us provide an example of SEM path diagram, which formula is:

$$y_1 = \gamma_{11} x_1 + \gamma_{12} x_2 + \zeta_1$$
$$y_2 = \beta_{21} y_1 + \gamma_{22} x_2 + \zeta_2 \quad (1.14)$$

Where:

y - endogenous variables,

x - exogenous variables,

γ - path coefficients for exogenous variables,

β - path coefficient for endogenous variables.

Here is quite simple example of SEM (1.14) with only two equations; also, this model is recursive, following the logic of equations order. Figure below (1.4) depicts the given SEM (1.14) in a graphical way:

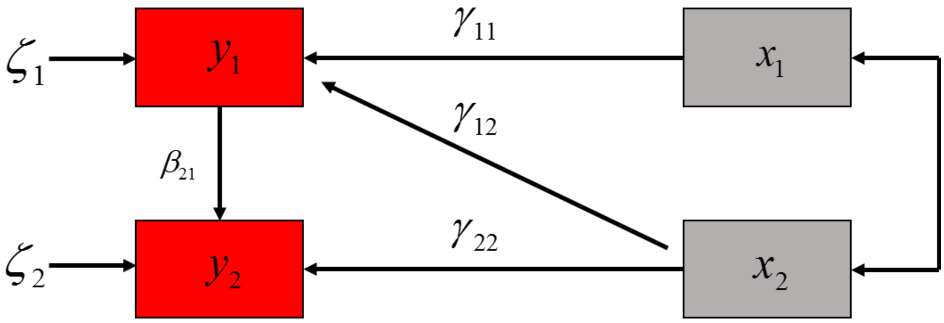


Figure 1.4 Path Diagram

Now, it is time to link all theoretical background about SEM to an implementation for residual operating income and discuss previous research work in this feald. Having interdependencies in operating income's drivers, there is a need of using SEM instead of just single equations. Relations between different accounting numbers are described by Tsal et al. (2008). Simultaneous equations were widely used before for valuation purposes. Nature of SEM is described by Wooldridge (2006), given with the meaning of simultaneity: it is "when one or more of the explanatory variables is jointly determined with the dependent variable" (p. 546). Recent fundamental research in company valuation, which could be counted as a framework, was done for a case of emerging markets (Ntim et al. 2015). This paper describes correlation between board size and firm's valuation, showing positive results. Earlier, sales forecasting with implementation of inventory and gross margin data was done for retail industry using SEM (Kesavan et al. 2010). Bergmann (2013) did her study in the similar way, but focusing on income forecasting instead of sales, and without constructing model for particular industry.

The framework, used by Bergmann (2013) and this masterwork, are suggested by textbooks about financial analysis (Penman, 2007; Lundholm and Sloan, 2007) – having interdependencies, forecasting process is not linear.

Bergmann (2013) in her model identified central value drivers: sales, operating expenditures and net operating assets. This value driver – endogenous variables – were modeled with a set of simultaneous equations. Aim of this masterwork is to complete author's model with industry specific exogenous variables.

Having a set of observations and SEM, there is not possible to use method of ordinary least squares (OLS). Having interdependencies, created regression will be biased and inconsistent. For this case, two-stage least squares method (2SLS) and three-stage least squares method (3SLS) are more suitable for obtaining values of parameters. Application of these methods into SEM are searched in details by McFadden (1999).

Chapter 2. Methodology

2.1 Estimation Model Development and Data Description

2.1.1 Application of Simultaneously Equations Model

The model in its first form will be align with model, proposed by Bergmann (2013). Author did a fundamental work, building models using simultaneous equations. She proved in this work existence of interdependencies between endogenous variables (SALES, OPEX and NOA) and its dependency on a set of exogenous variables. The amount of observations was very big – more than 19 thousands. The only one problem that this model covers very different industries. Having this too universal model, it is very doubtful to receive precise results for every industry. First step was precise determination the core of the model, proposed by Bergmann (2013). Author proposed four models and compare it using mean errors (median average percentage error– MdAPE; mean absolute error – MAE; root mean square error – RMSE). The most suitable model with less errors is the next:

$$\begin{aligned}
&SALES_{it}=\alpha_{10}+\alpha_{11}NOA_{it}+\alpha_{12}OPEX_{it}+\alpha_{13}SALES_{it-1}+\alpha_{14}OB_{it-1}+\\
&+\alpha_{15}FedRate_{it-1}++\alpha_{16}\Delta GDP_{t}+\alpha_{17}HI_{t-1}+\alpha_{18}CapIntI_{t-1}+\alpha_{19}MS_{it-1}+\\
&+\alpha_{110}CapIntF_{it-1}+\varepsilon_{it}\\
&OPEX_{it}=\alpha_{20}+\alpha_{21}SALES_{it}+\alpha_{22}OPEX_{it-1}+\alpha_{23}MS_{it-1}+\alpha_{24}\Delta PPI_{t}+\eta_{it}\\
&NOA_{it}=\alpha_{30}+\alpha_{31}SALES_{it}+\alpha_{32}NOA_{it-1}+\alpha_{33}FedRate_{t-1}+\upsilon_{it}
\end{aligned} \quad (2.1)$$

Where:

$SALES$ = net sales,
$OPEX$ = operating expenditures,
NOA = net operating assets,
OB = order backlog,
$FedRate$ = federal funds rate,
ΔGDP = change in gross domestic product,
HI = Herfindahl index,
$CapIntI$ = capital intensity of industry,
$CapIntF$ = capital intensity of firm,
MS = market share,
ΔPPI = change in product price index

Bergmann called this model (2.1) as a perfect foresight model. The estimation was done using three-stage least squares model (3SLS). Why exactly to use this model, not ordinary least squares (OLS)? Actually, in this case we have system of structural equations (in our case –

simultaneous), where we have endogenous variables (SALES, OPEX, NOA in our case) among explanatory ones. And, by having disturbance correlated with endogenous variables, we violate the main assumption of OLS model. Other word, estimating equations by OLS, we will force to do it separately, without taking into accounting interdependencies between them. Further, as we have explanatory variables among dependent in other equations, the error terms between equations will be correlated. Finally, results of our estimation will be biased. According to Davidson and MacKinnon (1993), 3SLS model uses an instrumental variable approach for estimates and generalized least squares (GLS) for taking into account correlation structure in the disturbances. Greene (2012) underlined three steps in 3SLS model:

- **First step**

  During this step, there is creation of instrumented values for all endogenous variables. The main purpose of these instrumented values is that they can be considered as predicted. This results from creating a regression of each endogenous variables on all exogenous. This step is crucial one for performing whole estimation and identical to one, which is using is two-least squares estimation (2SLS).

- **Second step**

  This step implicates an obtaining of consistent estimate for covariance matrix of equation disturbances. As a base for this step, model is using residuals from 2SLS estimation of each equation.

- **Third step**

  On this step, 3SLS model using estimates covariance matrix again, but now performing GLS-type estimation. Also, all endogenous variable on the right side of equations (which were determining) are replaced with instrumental variables.

Now, we can see a logic why this method is called “three-stage least squares”. Logic of choosing this method is quite understandable – we simply cannot perform good estimation model using OLS in our case. Actually, there are several methods for performing good unbiased estimations for structural equations. One of them is two-stages least squares (2SLS). This method is quite similar to 3SLS but have simpler estimation process: consists of two steps instead of three. 2SLS is the extension of OLS method. According to Angrist and Imbens (1995), we can derive principles of both steps of the model:

- **First step**

There is a creating of new variables in this step. This process is performing with a use of instrumental variables.

- **Second step**

In this step, there is an estimation of variables, created in previous step. Then, actual values are being replaced by new variable and model starts performing simple OLS estimation.

Usually, 3SLS model is more preferable for simultaneous equation than 2SLS. Generally, reasons for choosing model estimation is up to researcher, but sometimes, there are some external limiters for models using. For example, in this masterwork statistical analysis was done in program environment Stata 12, where usage of 3SLS estimation is limited. Nevertheless, first step of performing technical analysis was to derive core model from model, proposed by Bergmann. This step is necessary, because proposed foresight model contain many not-industry specific variables, which will affect precise estimation for our chosen industry.

The core model, taken as a base for further analysis has the next form:

$$\begin{aligned} SALES_{it} &= \alpha_{10} + \alpha_{11} NOA_{it} + \alpha_{12} OPEX_{it} + \alpha_{13} SALES_{it-1} + \varepsilon_{it} \\ OPEX_{it} &= \alpha_{20} + \alpha_{21} SALES_{it} + \alpha_{22} OPEX_{it-1} + \eta_{it} \\ NOA_{it} &= \alpha_{30} + \alpha_{31} SALES_{it} + \alpha_{32} NOA_{it-1} + \upsilon_{it} \end{aligned} \quad , \qquad (2.2)$$

Where:

$SALES$ = net sales,

$OPEX$ = operating expenditures,

NOA = net operating assets.

The one independent variable, which could be used for further analysis is order backlog (*OB*). This is firm-specific indicator, which, probably, has a huge influence on a sales forecasting (Elliot and Uphoff, 1972). Order backlog is the sales order waiting to be filled, actually, a negative component, whom all companies want to avoid. Unfortunately, as we a focusing current study on Russian industry, there is no consistent information about order backlog. That probably could be improved in a future with developing of data disclosure politics. Nevertheless, for now it is just a possible way for future improvement.

For achievement of set goals, we need to do precise estimation and to find a between all set variables (SALES, OPEX and NOA) and their interdependency. Actually, later we will be adding another independent industry-specific and firm-specific variables for more precise estimation. To find these interdependencies and influences, we will perform different regressions and make regression analysis. Therefore, first, we need to define statement “regression analysis” and understand the process behind it.

Regression analysis is a statistical process for estimation relationships between variables. It used for statistical models – sets of taken assumptions and received data. So, in regression analysis, we built and estimate regressions – approaches for relationship modelling. All regression can be depicted in the next form:

$$Y = f(X, \beta) \quad (2.3)$$

Where:
Y - dependent variable,
X - independent variable,
β - unknown parameter.

In more simple form, regression is a function of Y on X, and the main purpose is to estimate unknown variable β. What kind of function and how to estimate these unknown variables are the most important research questions. There are many different types of regressions for different purposes and we will not stop an each of them. Main types of regression and logic of estimation will be explained further, during the current regression analysis.

Now we have an understanding, with which kind of data we will be operating and in next part descriptive statistics will be described. But, first, we need to pay attention on statistical software for performing regression analysis.

Usually, the choice of statistical program is up to researcher. In this masterwork, main criteria were: friendly interface and availability for usage. Under these criteria, Stata[2] was chosen as a main tool for regression analysis. This package was created in 1985 by StataCorp and remain one of the most popular tool of general-purpose statistical analysis. Table below illustrates versions of data and capabilities of this software package:

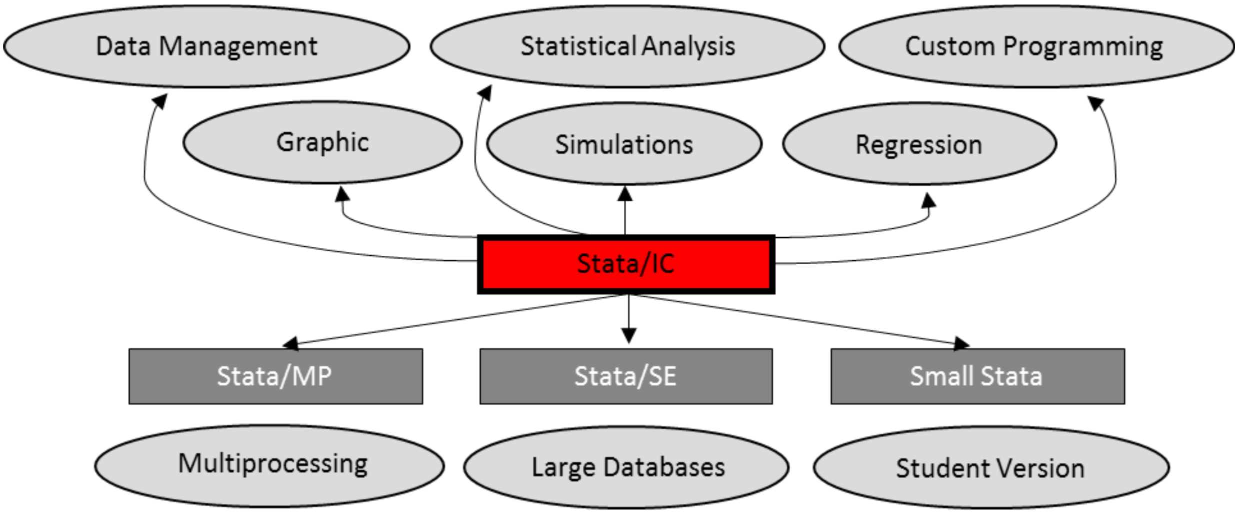


Figure 2.1 Stata Package version and capabilities

2 http://www.stata.com/ Stata Package Official website

2.1.2 Data and Descriptive Statistics

Oil and gas exporting were driving Russian economy for many years. Exactly these sources of energy helped Russia to recover from depression in 1999.

Three main issues arise from an oil policy inside the country:

1) Natural characteristics of oil and its reserves
2) Political governance
3) Technical and economic sides of oil exploration

This purpose of this master work is to dig into the third issue.

Even our national currency depends on world prices of crude oil. Influence of oil exporting on monetary policy is overviewed by Russian ex-minister of finance (Kudrin, 2014).

Nevertheless, this work is not going to overestimate the value of this industry to Russia. In many news sources, it is written that contribution of oil industry into federal budget for year 2105 was more than 43%[3]. We need to distinguish federal and consolidated (including regions) budgets. Oil industry contribution into consolidated budget is less than 20%[4].

Second reason for creating valuation model for oil sector in Russia is a fact that our country has a seventh place in amount of proven oil reserves.[5] Only six countries: Venezuela, Saudi Arabia, Canada, Iran, Iraq and Kuwait have bigger oil reserves. However, according to the fact that oil is a strategic raw, information about reserves could be distorted by some countries.

In previous year, there was extracted 532.9 million tons of oil (including gas condensate) in Russia.[6] In comparison with year 2014, oil extraction had increased by 1.3%. Amount of exported oil is 244.5 million tons, which is 9.4% more than in year 2014. Share of imported oil is 45.9%.[7] In 2015, weight of exported oil of all exported goods was 26.1%, in 2014 – 30.9%. Almost 300 companies in Russia have government permission for oil extracting, having more than 30% as vertical integrated companies.

Having such sharp decline in oil prices, there is a hard situation for all Russian energy market. Cash inflow from oil exporting was 87 billion dollars, much less than in 2014, it was almost 150 billion dollars.

The perspectives of Russian oil industry depends on world prices on crude oil. Moreover, Russia needs to be aware of movements of main competitors – Saudi Arabia and USA. But oil prices cannot be the only crucial factor in oil extracting policy. Different regions have different costs of oil extracting – from 4 dollars in Saudi Arabia to 32 dollars for American slate oil. For recently opened oilfield in Russia, costs are 16 dollars per barrel, for old oilfields – just 6.

---

3 http://ria.ru/infografika/20130912/958932396.html

4 http://info.minfin.ru/kons_doh_isp.php

5 https://www.cia.gov/library/publications/the-world-factbook/rankorder/2244rank.html

6 http://www.vedomosti.ru/business/articles/2016/02/03/626546-dobicha-nefti-rossii-rastet-blagodarya-bashnefti-gazprom-nefti-tatnefti

7 http://www.gks.ru/bgd/free/b04_03/IssWWW.exe/Stg/d06/38.htm

Many researches were dedicated specifically to forecasting of Russian oil industry.

Sample comprises top Russian oil companies. Search was done using sources, available in library of Graduate School of Management. First step was downloading all necessary information using Thomson Reuters Datastream. Then, due to not large sample size, to avoid any elimination of omitted variables, all variable very checked manually with Thomson Reuters Eikon. Because of working with top Russian oil companies, more detailed information about some missing or questionable variables were checked in SCREEN and SPARK. All variables measured in dollars – there is not possible precisely measure firm's performance in rubles due to high correlation between crude oil price and ruble's appreciation/depreciation. All data is taken for period last 15 years. From year 2000 until year 2014, which information was available for most companies for the moment of writing this masterwork. This period includes increasing price of oil from year 2000, period of world economic crisis 2007-2009 and crisis in Russia in 2014. The best conditions for testing the model.

Here is short review of eight Russian oil companies, which variables were taken and used in current research:

- *Bashneft*. One of the oldest Russian oil companies, established in 1932. Has 170 oil plants in three oil extraction regions in Russia – Volga-Urals province, Timan-Pechora and Western Siberia. Oil production is about 17 million tons per year. For year 2014 has 3,644 million barrels of 3P (proved + probable + possible) reserves. Classification of oil reserves type will be given later. 3.4% share in oil production in Russia. Total revenues for December, 31, 2014 were 637,231 million of Russian rubles with operating profit of 76,997 million of rubles.[8]
- *RN Holding*. Formerly TNK BP, company was acquired by Rosneft in 2013. Previously one of the leading Russian oil companies, third place in amount of oil extraction with a share of 15% in Russia. Had been operating in Western and Eastern Siberia, Far East and Volga-Ural region. Turnover for last functioning year was 688 billion of Russian rubles with operating profit of 202.5 billion rubles.[9]
- *SurgutNeftegaz*. Third biggest company in amount of oil production – 61 million tons and 12% share in Russia. Company operates 22,380 wells and has a class C1 77 million tonnes of recoverable oil reserves. Company's operations predominantly done in Western Siberia, having 62 out of 67 oil field there. For year 2014 company's total sales were 863 billion of Russian rubles with pretax profit of 1,068 million rubles.[10]

8 http://www.bashneft.com/ Official website

9 http://www.finam.ru/ Article 2014 about RN Holding's performance

10 http://www.surgutneftegas.ru/ Official website

- *Slavneft*. Russian oil company. Has a eighth place in oil extraction amount with a share of 2.9%. Operates on 31 oil fields in Khanty-Mansi Autonomous Area and Krasnoyarsk region. In year 2014 had 7,831 million of 3P oil reserves. Oil production is 16.2 million tons of crude oil. Total turnover was 10,334 million of Russian rubles with total loss on sales of 807 million rubles.
- *Rosneft*. The largest world company in oil production and amount of oil reserves. In year 2014 oil extraction was 204.9 million of tons. Share in Russia is 37%. Number of C1 + C2 hydrocarbon reserves is 129 million of barrels. Rosneft operates in Western Siberia, Eastern Siberia, Timan-Pechora, Central Russia, southern part of European Russia and the Russian Far East. Sales were 5,503 billion of rubles wuth EBITDA of 1,057 billion of rubles in 2014.
- *Lukoil*. One of the largest Russian company with revenues exceeds 10.000 billion rubles and net income of 350 billion rubles. Company has a 16.4% share in Russian oil extraction amount and 2% share of world oil extraction. Lukoil operates in 559 oil plants and has 27.8 billion barrels of 3P oil reserves. Main company's operations are in Western Siberia, Ural and Timan-Pechora regions.
- *Tarneft*. Russian oil company, which primarily operates in European part of Russia. Company has 851.5 million barrels of proved oil reserves and produce 26 million tons of oil annually. Main crude oil amount comes from six oil fields. Turnover in 2014 was 392 billion rubles with net income of 82 billion rubles.

The list includes the eight biggest Russian oil companies. The main financial variables were taken from Thomson Reuters Eikon, in same accounting standard – IFRS.

It was mentioned above that we operates with data in statistical models, so we need to describe different types of data and to use it for regression analysis. We will not dig deep into specifications of quantitative or qualitative data and highlight only three main types of data we will be interested in current research work.

- **Cross-sectional data**

This type of data means cross section of a study population. It can be collected by observing many different subjects – in our case, for example, firms. The main principal is to take different subjects and to compare their differences at one period of time, of without taking into account any differences in time periods.

- **Time-series data**

Different type of data, in which accent is taken for measure time points. Sample usually is just one subject, small-scale or aggregate data taken in different periods of time. Purpose of this type of data is to find differences not between subjects, but between their measures in different points of time.

- **Panel data**

Also called time-series cross-sectional data. This type of data is a combination of previously described methods. It is a multidimensional data involving measures over time. Here we obtain data about different subjects, their measurements over time and look at both parameters simultaneously.

As we described in introductory part, main formula of residual operating income we will be working with is following:

$$ReOI_t = (SALES_t - OPEX_t) \cdot (1 - tax) - WACC \cdot NOA_{t-1}, \tag{2.4}$$

Where:

$SALES$ - total revenues,

$OPEX$ - operational expenditures,

$WACC$ - weighted average cost of capital,

NOA - net operating assets.

Endogenous variables (EV) are determined inside the model system. They will be chosen as a main value drivers. We discussed those three main value drivers previously: SALES, OPEX and NOA. As we can see, all those variables appear in the formula (2.4). One more variable is WACC, which will be taken separately and will not be included in simultaneous equations model. Let us give precise determination of an each variable.

- **Sales**

Actually, sales forecasting is the basic step of almost all forecasting models. This variable often determines, how company will be growing and developing in the future. The main weak point of forecasting sales is that often this step is only one done for forecasting process (Koller, 2010). Main value drivers exactly for sales are competition level, product substitutes, possible patents and any positive expectations, associated with brand (Penman, 2010). Actually, amount of factors, which influence sales is much wider, so, often, they are just not included into forecasting assumptions (Richardson, 2010). One possible factor influence sales is interdependency with operational expenditures and net operating assets. **The purpose of this masterwork, once again, is to create disaggregated residual operating income forecast, taking into accounting possible interdependencies between independent variables**. Sales amount are taken directly from databases like Thomson Reuters Database without any additional calculations.

- **Operating expenditures**

Knowing sales given level, operating expenditures determines the level of profit margin. Usually, profit margin is driven by economy of scale, consummation of technology process, competiveness in supplier, labor sources, and so on (Penman, 2010). Koller (2010) recommends forecasting of OPEX in line according to SALES level. So, more SALES, more OPEX and otherwise. Usually, this model should work, logically. However, this is truth not for all cases, OPEX level can easily have other pace, than SALES, depending on market conditions. OPEX cannot be taken directly and needs to be calculated. The formula for deriving this variable is next:

$$OPEX = COGS + SG\&A + D\&A, \tag{2.5}$$

Where:
$OPEX$ - operational expenditures,
$COGS$ - cost of goods sold,
$SG\&A$ - selling, general and administrative expenses,
$D\&A$ - depreciation and amortization.
In this masterwork, we will be looking at OPEX of biggest Russian oil companies. According to Cheng (2005), big market share leads to better manifestation of economy of scale and more efficient production.

- **Net operating assets**

With determining level of $OPEX$ and knowing level of $SALES$, we estimate level of profit margin. Profit margin is a measure of profitability. In contrast to profit margin, asset turnover measures a production efficiency. It measures potential of assets to boost and generate sales. Prediction of sales amount will help us in determining of future net operating assets amount only indirectly, we need to take into account future changing of working capital and long-term assets. There is a huge chain of dependent value drivers. For example, long-term assets are determined by property, plant and equipment ($PP\&E$), which are driven, in its turn, by implementation of new technology and new products creation. The same logic with determining drivers of working capital by splitting it into accounts payable, inventories and accounts receivable. At the same time, net, working capital and property, plant and equipment can be forecast depending on sales amount, and what is why we again can implement rude assumption of taking net operating assets as percentage of sales. And, again, we found aa evidence of interdependency and that is why we will support an implementation of simultaneous equations model. But first, we need to take a data about net operating assets in different companies. Again, as in a case with operating expenditures, this parameter could not be taken directly. This variable will be calculated in accordance with a set of studies (Nissim and Penman, 2001; Callen and Segal, 2005; Soliman, 2008). Net operating assets are defined as operating assets (*OA*) minus operating liabilities (*OL*). So, we have a following formula:

$$NOA = OA - OL, \tag{2.6}$$

Where:
NOA - net operating assets,
OA - operating assets,
OL - operating liabilities.
At the same time, we need to define operating assets and operating liabilities:

$$OA = TA - FA, \tag{2.7}$$

$$FA = CSTI + OIA, \tag{2.8}$$

Where: OA - operating assets,
TA - total assets,
FA - financial assets,
$CSTI$ - cash & short-term investments,
OIA - other investments and advances.

$$OL = TL - FL, \tag{2.9}$$

$$FL = LTD + DCL + MI, \tag{2.10}$$

Where: OL - operating liabilities,
TL - total liabilities,
FL - financial liabilities,
LTD - long-term debt,
DCL - debt in current liabilities,
MI - minority interest.
By combining all above, we can define net operating assets as following:

$$NOA = TA - CSTI - OIA - TL + LTD + DCL + MI. \tag{2.11}$$

We defined all three main endogenous variables. Next part will be devoted to different ways of result estimation.

2.1.3 Application Three-Stage Least Squares Initial Estimation.

As for now, the main priority is to check the usage of this core model for Russian oil industry. Without this step, we cannot process with further analysis. For these purposes, this model were estimated as in underlying article (Bergmann, 2013) with a use of a 3SLS estimation. Our sample consists of an eight companies for 15 years. Having some missed variable, in total, we operated with 106 observation. Having multidimensional data, our data has a panel, or longitudinal, type. However, for estimating system of equations (2.2), we have a right to look only at a company-year observation and use 3SLS model, without paying a lot attention now for determining each observation's company or year.

Received coefficients are presented in the table (2.1) below:

Table 2.1 Three-stage least squares estimation

	Equation (1) $SALES_{it}$ Coef. [Std. Errors]	Equation (2) $OPEX_{it}$ Coef. [Std. Errors]	Equation (3) NOA_{it} Coef. [Std. Errors]
$SALES_{it}$		0.3066*** [0.1431]	-0.1895 [0.4664]
$OPEX_{it}$	-0.0782 [0.2727]		
NOA_{it}	0.0029 [0.0065]		
$SALES_{it-1}$	1.2419*** [0.2342]		
$OPEX_{it-1}$		0.7327*** [0.2200]	
NOA_{it-1}			0.8738*** [0.0554]
R^2	0.8424	0.9337	0.7108
*p<0.1, **p<0.5, ***p<0.01			

While estimating system of equations (2. 2), were received a P-value < 0.0000 for all three equations, which means, we have significant equation and can use them for further analysis. From the equation (1), we can see that only SALES amount in previous year can have some influence on SALES in present period. Having positive coefficient and its number of 1.2419 with probability of mistake is less than 5 %, this result is quite expectable. Coefficient for OPEX in present time is negative, which can be explained by economy of scale, but it is not significant having that big standard error and big probability of mistake. The same situation with net operating assets – seemingly no influence of net operating assets on sales amount in present year. R-square for equation (1) is quite big (0.8424), so regression approximates real data point quite well.

Result for operation expenditures were more expected. As in Bergmann's (2013) model, we receive a high level of significance for sales coefficient (0.3066 with a $p < 0.05$). The positive sign could be explained as if we have more sales, we need (or, in some cases, may) spent more money for operating expenditures. The link between operating expenditures in present and past year confirmed. Coefficient for OPEX in previous year is 0.7327 with high level of significance ($p < 0.05$). Having coefficient of determination equals 0.9337, we can say, that, generally, regression fits the given data very good.

Last equation (3) is dedicated to finding possible dependency of net operating assets in present year on sales level in present year and net operating assets in year $t-1$. Actually, by performing estimation, only NOA_{it-1} coefficient was found as significant (0.8738, $p < 0.05$). The coefficient number close to one tells us about close net operating assets year-to-year amount. Determination coefficient is 0.7108, so goodness of fit of current regression is normal one.

The reason, why some coefficients became insignificant is that we have incomplete model with few independent variable. Having SALES, OPEX and NOA as dependent and determining variables simultaneously will not permit us to receive precise results. Nevertheless, by performing this step, we have found that simultaneous equations model is applicable for Russian oil industry.

Next step will be dedicated to finding specific exogenous variables.

2.1.4 Extended model

For this moment, we are on way of deriving o new model from core one. This model will not be completely new – it just contains some improvements for residual operating income model, taking into account interdependency between endogenous variables.

Now, we need to decide, which exogenous variables we need to include into model. Here is list of variables with detailed explanations, why do we need to include them into model and influence on which parameter we expect to receive.

- **Oil price**

By oil price, we mean price of Brent Crude, which is a benchmark for oil pricing worldwide. This parameter should obviously influence amount of sales for oil companies. It is quite obviously, if we will look on one of fundamental economic equations:

$$R = P \cdot Q, \tag{2.12}$$

Where:

R - revenues or sales,

P - price,
Q - quantity.

Fundamental implementation of oil price into econometric model was done by Osmundsen et al. (2006). Although, the main purpose of research was to define most appropriate measure for oil companies performance: return on average capital employed (RoACE) or enterprise value to debt-adjusted cash flow (EV/DACF), the significant influence of oil prices were strengthen. The model, proposed by author, has the next form:

$$EV/DACF = A_i + \alpha P_t + \sum_{i=1}^{n} \beta_i KPI_{it} + \gamma R_{it} + u_t, \tag{2.13}$$

Where:
EV - enterprise value,
$DACF$ - debt-adjusted cash flow,
P - crude oil price,
KPI - vector of key performance indicators,
R - return on average capital employed.

This model is given not for interpretation purposes, but as an example of fundamental work findings about its significant influence. Oil prices were taken from Yahoo Finance[11] on a daily basis for more precise results calculation. Then, annual average price were calculated. As we are working with annul data for all other variables, by take average annual oil price, we ignore sharp changes in oil prices movements. Nevertheless, the whole trend of oil pricing is taken into account.

- **Oil price volatility**

That means oil price volatility? Volatility is measure of statistical dispersion on some given object (usually, market security). In our case, it is a low responsiveness or "inelasticity" of both demanders and suppliers of quick price adjusting in a short-term. The most recent work done in this field was dedicated to study oil price volatility as a main factor for oil speculation (Beidas-Strom and Pescatori, 2014). By doing a VAR (vector autoregression) estimation, author found a significant influence of oil price volatility on oil supply/demand equilibrium changes. Damodaran (2005) in his book “Investment Valuation” made a fundamental analysis with a set of commodities companies with a next implementation “value of a natural resource company depends not just on the price of the natural resource but also on the expected volatility in that price”. That is why we decided to include this parameter into our

11 http://finance.yahoo.com/ Official website

model. Actual historical oil price volatility was calculated directly from given oil prices, without searching it on external sources. By knowing daily prices and its daily logarithmic return, we may calculate annual volatility by formula:

$$\sigma = \frac{\sigma_{SD}}{\sqrt{P}}, \qquad (2.14)$$

Where:

σ - volatility,

σ_{SD} - standard deviation,

P - time period (day).

For calculating annual volatility, we need to add into formula 252 trading days in one year:

$$\sigma_{ANNUAL} = \sigma_{SD} \cdot \sqrt{252}, \qquad (2.15)$$

Having calculated annual volatility for 15 periods, we will input it into our model.

- **Oil Production**

In recent study "Oil and gas production handbook" (Devold, 2013), focus was done on upstream part of oil production and on one the most important drivers – amount of oil produced per year. As we are working with Russian oil industry, we cannot ignore this factor as Russia is one of the world's leaders in amount of oil produced. For the end of 2015 situation was next:

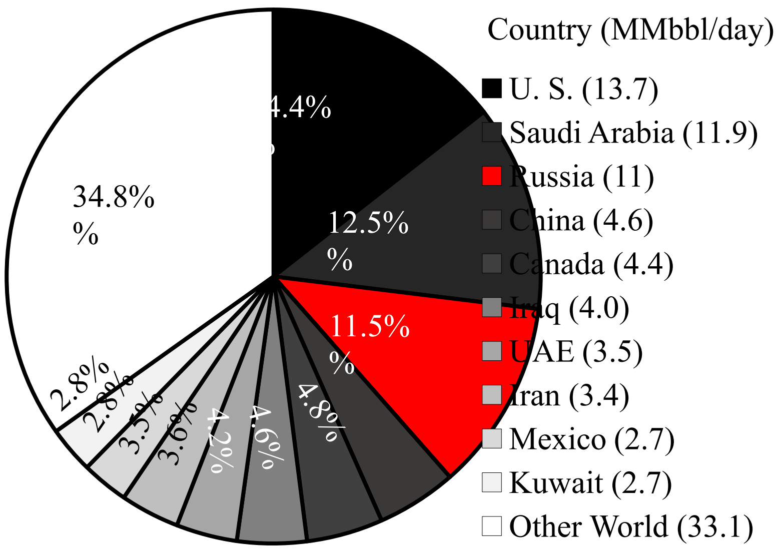


Fig. 2.2 Daily crude oil production[12]

12 Source: US Energy Information Administration

So, oil production is very important indicator of company's performance, because it has a high influence at least on sales amount. This influence was underlined by Gurvich (2004) in his macroeconomical valuation of Russian oil sector. Number of oil production amount mostly available in companies' annual reports. Sometimes it is given in cubic metres. Conversion factor taken the same as set by New York Mercantile Exchange (NYMEX)[13] – 6.29287 barrels per cubic metre.

- **Oil Reserves**

Oil reserves is the current amount of technically available oil. Next very important parameter for precise companies' performance calculation. Then we are talking about oil reserves, we need to determine their types:

1) Proved reserves. Classification as P1. Those reserves has a level of confidence of 90% that this reserve oil can be extracted and used (Wright and Gallun, 2008). Additional division if for proved developed reserves (PD) and proved undeveloped reserves (PUD). First type reserves can be extracted with existing technical equipment and wells or with minimal investments. Second type require additional capital raising.

2) Probable reserves. In the oil industry proved + probable reserves referred as 2P. Those reserves claimed as 50% confidence as being extracted.

3) Possible reserves. Proved + probable + possible - 3P classification. Has only a 10% confidence level due to geological interpretation or current absent of technology.

Since year 2010, Security Exchange Commission (SEC)[14] allowed companies to publish information about all types of reserves in their reports. Previously, only information about proven reserves was allowed for publishing. In Russia, categories A, B and C1 refer to proved reserves. Category C2 includes both probable and possible reserves[15]. Russia has a huge oil reserves, but not the leading position in world:

13 http://www.cmegroup.com/company/nymex.html NYMEX official website
14 http://www.sec.gov/ Security Exchange Commission, official website
15 Source: Society of Petroleum Engineers

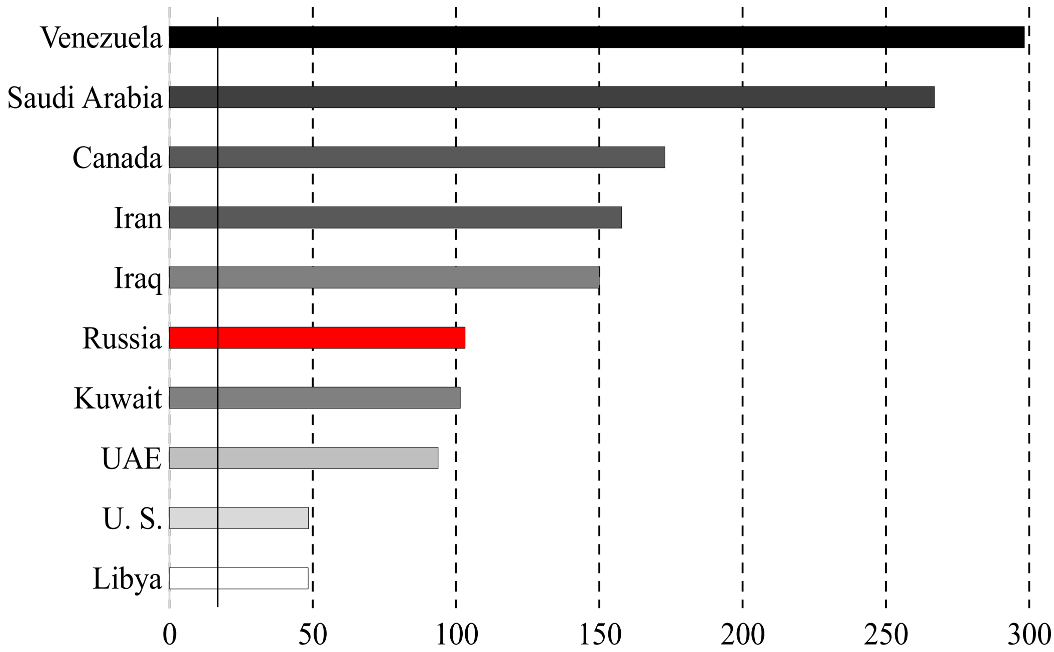


Fig. 2.3 World proven oil reserves by country, BBL[16]

Oil reserves is the most difficult to find data, because Russian companies are not obliged to publish this information in their reports. Nevertheless, information was found and presented for further research.

- **EV/EBITDA and EBITDA/barrel ratios**

As for the last independent variables, decision was to take some industry ratios. First, it is very popular one for different industries: enterprise value to earnings before interest, taxes, depreciation and amortization. Second, it is more industry-specific ration - earnings before interest, taxes, depreciation and amortization to annual oil production in barrels. For more precise understanding why we need to use these ratios, we derive formulas for each part of rations. Enterprise value is an economic measure of a company in according to its market value. This metric do not take into account capital structure of a company. That is why, if we are comparing company with a very different level of debt, this ratio can be more precise than, for example, Price/Earnings ratio. The general formula of EV is:

$$EV = CSE + D + MI + PE - C\&CE\,, \tag{2.16}$$

Where:

EV - enterprise value,

CSE - market value common shareholders' equity,

D - debt at market value,

16 Billions of barrels, source: US Energy Information Administration

MI - minority interest,
PE - market value of preferred equity,
$C\&CE$ - cash and cash equivalents.

In denominator of this ratio there is an EBITDA. Formula can be written as next, in according with our main variables:

$$EBITDA = SALES - OPEX + D\&A, \tag{2.17}$$

Where:
$EBITDA$ - earnings before interest, taxes, depreciation and amortization,
$SALES$ - total sales or revenues,
$OPEX$ - operational expenditures,
$D\&A$ - depreciation and amortization.

Ratio EV/EBITDA is used as the most important for valuing capital-intensive companies with a high level of depreciation and amortization. Second ratio is more specific for oil industry and uses in denominator amount of crude oil in barrels produced annually. This metric is widely used by oil companies, for example, by Rosneft.

We discussed six main components, which will be included into the model as independent variables. The table below presents the short summary about all variables and their probable influence on SALES, OPEX or NOA:

Table 2.2 Exogenous variables.

Variable	Description	Expected sign
Macroeconomic factors		
Oil Price	Oil price for a year at time t. Taken as annual average value. Measured in dollars.	+ (SALES) - (OPEX)
Oil Price Volatility	Oil price volatility for current year. Calculated on historical basis. Taken as average annual volatility at time t.	- (SALES) + (OPEX)
Company-specific factors		
Oil Production	Annual oil production in barrels for year at time t.	+ (SALES) - (OPEX)
Oil Reserves	Proved oil reserves in barrels for year at time t.	+ (SALES) + (NOA)
EV/EBITDA	Ratio, enterprise value of company to earnings before interest, taxes, depreciation and amortization for year at time t.	- (SALES) + (OPEX)
EBITDA/Barrel	Ratio, earnings before interest, taxes, depreciation and amortization to annual oil production in barrels for year at time t.	+ (SALES)

On this stage, we are interested in interpretation of interdependency of endogenous variables and influence of new included variables, not in forecasting process. So, all variables will be taken in present time t.

2.1.5 Generalized Two-Stage Least Squares Estimation

Now, there is a question of which estimation model to choose. The first estimation was executed with a help of 3SLS model with a rough estimation of just taking company-years observation, without paying attention on panel data. Now, we cannot perform estimation this way. The main problem is that we have now not only endogenous variables, which are unique for each company in each year, but some exogenous variables, which are the same for all companies in one year. 3SLS model is not working with panel data, at least in Stata environment. The choice was done on G2SLS model (Balestra and Varadharajan-Krishnakumar, 1987) – this a model uses instrumental variables and 2SLS estimation for panel data. This models offer five different estimators, which are called models with different effects. When we are speaking about panel data estimation, we are usually mean two types of models:

- **Fixed Effects Model**

  It is a statistical model, where observed quantities are not random. The fixed effects assumption is that individual effects are correlated with the independent variables. If random effect assumption holds, the fixed effects model are less efficient, but when it does not holds, we may use only fixed effects model.

- **Random Effects Model**

  Also known a variance components model. This is a type of hierarchical linear model, whose difference of population relates to its hierarchy. In our case, random effects model is used for panel data with assumption of no fixed effects.

Actually. There are several more models like Between-Effects model (BE) or First-Difference model (FD), but they are not useful in this masterwork. Here, or choice will be done between Fixed Effects model and Random Effects model. For determination, which model to choose, usually Durbin–Wu–Hausman (also known as Hausman specification test) is used (Hausman, 1978). This tests evaluates consistency of estimators, under null hypothesis is Random Effects, under the alternative – Fixed Effects model.

After running the regressions using Fixed Effects model and Random Effects model (there is an algorithm of performing this test), Hausman showed us that the more consistent variant will be to use Random Effects model. Actually, there is one more implementation of 2SLS with Random Effects, except G2SLS – EC2SLS (Baltagi, 2013). But second model is not so popular because of its specification.

After describing and including exogenous variables, our model will have next form:

$$\begin{aligned}
&SALES_{it}=\alpha_{11}OPEX_{it}+\alpha_{12}NOA_{it}+\alpha_{13}SALES_{it-1}+\alpha_{14}Oilprice_{t}+\\
&+\alpha_{15}Oilvolatility_{t}+\alpha_{16}Oilreserves_{t-1}+\alpha_{17}Oilproduction_{t-1}+\\
&+\alpha_{18}EV/EBITDA_{t-1}+\alpha_{20}EBITDA/Barrel+\varepsilon_{it}\\
&OPEX_{it}=\alpha_{21}SALES_{it}+\alpha_{22}OPEX_{it-1}+\alpha_{23}Oilprice_{t}+\alpha_{23}Oilvolatility_{t}+\\
&+\alpha_{18}EV/EBITDA_{t-1}+\eta_{it}\\
&NOA_{it}=\alpha_{31}SALES_{it}+\alpha_{32}NOA_{it-1}+\alpha_{33}Oilreserves_{t-1}+\upsilon_{it}
\end{aligned} \quad (2.18)$$

Where:

$SALES$ - total sales or revenues,

$OPEX$ - operational expenditures,

NOA - net operating assets,

$Oilprice$ - average annual price for barrel of crude oil,

$Oilvolatility$ - average annual volatility for price of barrel of crude oil,

$Oireserves$ - company's oil reserves,

$Oilproduction$ - company's annual oil production,

$EV/EBITDA$ - enterprise value to earnings before interest, taxes, depreciation and amortization.

$EBITDA/Barrel$ - earnings before interest, taxes, depreciation and amortization to oil annual production in barrels.

By performing G2SLS estimation, received coefficients are next:

Table 2.3 Generalized two-stage least squares estimation results

	Equation (1) $SALES_{it}$ Coef. [Std. Errors]	Equation (2) $OPEX_{it}$ Coef. [Std. Errors]	Equation (3) NOA_{it} Coef. [Std. Errors]
$SALES_{it}$		0.8629*** [0.0329]	-8.8891* [6.4647]
$OPEX_{it}$	1.1340* [0.0363]		
NOA_{it}	0.0045*** [0.0032]		
$SALES_{it-1}$	0.0204 [0.0355]		
$OPEX_{it-1}$		0.0448 [0.0474]	
NOA_{it-1}			0.7298*** [0.0985]
$Oilprice$	4.2298*** [17.0101]	-5.7582*** [15.9715]	
$Oilvolatility$	-392.5103 [3380.9874]	-1164.238 [3938.259]	
$Oilreserves$	-0.3466*** [0.2281]		5.0383 [4.8324]
$Oilproduction$	14.5890*** [3.8557]		

$EV / EDBITDA$	-199.8089 [166.8622]	263.0131 [180.4787]	
$EBITDA / Barrel$	21.7232 [45.7983]		
R^2	0.9823	0.9850	0.7039
*p<0.1, **p<0.5, ***p<0.01			

After estimating system of equations (2.16), were received a P-value < 0.0000 for all three equations, which means, we have significant equation and can use them for further analysis. From the equation (1), we can see that only part of coefficients become significant for SALES influence. Coefficient of OPEX at time *t* is significant, but only with confidence level of 10%, sign is positive. Coefficient for NOA at time *t* is unexpectedly small – 0.0045 with a high level of significance (p<0.01). In contrast with initial model (2.2), by implementing extra exogenous variables, SALES influence at time *t-1* become insignificant. This fact was quite surprising. As it was expected, oil price has a high significance level of influence on sales. As we can see, from the result of regression analysis, coefficient of oil volatility influence on SALES is not significant. Both oil production and oil reserves showed a significant influence on SALES with coefficients 14.5890 and 0.3466 respectively. Both coefficients for EV/EBITDA and EBITDA/barrel seems to be insignificant. R-square for equation (1) is quite big (0.9823), so regression approximates real data point quite well.

In equation (2) coefficient for SALES is significant (p<0.01) with value of 0.8629. Again, in contrast with system of equations (2.2), endogenous variable at time *t-1* is not significant (in our case - $OPEX_{it-1}$). Oil price coefficient is significant with value of -5.7582. Negative sign follows the logic that with higher price OPEX will be less. In equation (2) as in equation (1), coefficients for oil price volatility and ration EV/EBITDA become insignificant. Having coefficient of determination equals 0.9337, we can say, that, generally, regression fits the given data very good.

Last equation (3) is dedicated to finding possible dependency of net operating assets. Actually, by performing estimation, two out of three coefficients was found as significant. Coefficient of SALES at time *t* is found as significant, but at confidence level of only 10%. High level of confidence showed a dependency on NOA at time *t-1* (0.7298, p<0.01). Determination coefficient is 0.7108, so goodness of fit of current regression is normal one.

By performing G2SLS regression, now we can see a set of exogenous variables investor should pay attention while performing valuation in Residual Income Operating Model. As it was found, investor should carefully estimate oil price, oil production and reserves, not only predict SALES amount. But for now, it very complicated to take into account all these variables and their interdependencies, so next part will be dedicating to building forecasting model based on this SEM model, studied in chapter 2.

2.2 Practical implications

2.2.1 Application of Panel Vector Autoregression

For further performing of forecast analysis, we need first to define correct method. Simultaneously taking interdependency of endogenous variables and lagged exogenous variables will be technically difficult. Under these conditions, taking G2SLS estimation is not beneficial. In this masterwork, choice was done on panel vector autoregression models (PVAR). First time-series vector autoregression model was introduced in literature as useful alternative of multivariate simultaneously equations models (Sims, 1980), which we used in second chapter. The main feature of this model is that all variables is treated as endogenous, although some restrictions exists (Abrigo and Love, 2015). From the very beginning of program development, PVAR models did not work with panel data. It found huge implementation in different fields and researches after adjusting settings for working with panel data (Holtz-Eakin, Newey and Rosen, 1988).

First usage of PVAR in Stata was described in details by Love and Zicchino (2006). Actually, this package is not in-built in Stata, so there is a necessity of downloading of this extra-pack. For performing estimation by PVAR and receive forecasting model, we need to take lagged variables from all variables (endogenous and exogenous). Which variables to include was decided in previous chapter – only those, which have a significant influence on SALES, OPEX and NOA. Received model should have the next form:

$$\begin{aligned}
&SALES_{it}=\alpha_{10}+\alpha_{11}SALES_{it-1}+\alpha_{12}OPEX_{it-1}+\alpha_{13}NOA_{it-1}+\alpha_{14}Oilprice_{it}+\\
&+\alpha_{15}Oilproduction_{it-1}+\alpha_{16}Oilreserves_{it-1}+\varepsilon_{it}\\
&OPEX_{it}=\alpha_{20}+\alpha_{21}SALES_{it-1}+\alpha_{22}OPEX_{it-1}+\alpha_{23}NOA_{it-1}+\alpha_{24}Oilprice_{it}+\\
&+\alpha_{25}Oilproduction_{it-1}+\alpha_{26}Oilreserves_{it-1}+\eta_{it}\\
&NOA_{it}=\alpha_{30}+\alpha_{31}SALES_{it-1}+\alpha_{32}OPEX_{it-1}+\alpha_{33}NOA_{it-1}+\alpha_{34}Oilprice_{it}+\\
&+\alpha_{35}Oilproduction_{it-1}+\alpha_{36}Oilreserves_{it-1}+\upsilon_{it}
\end{aligned} \tag{2.19}$$

Where:

$SALES$ - total sales or revenues,

$OPEX$ - operational expenditures,

NOA - net operating assets,

$Oilprice$ - average annual price for barrel of crude oil,

$Oireserves$ - company's oil reserves,

$Oilproduction$ - company's annual oil production.

As you can see from formula (2.19), oil price is taken at time *t*. The reason is that oil price is too volatile and its implementation just ruins the models. Experimentally were proved that after inputting of lagged oil price, all coefficients become insignificant. From this

assumption, we have one very important notification about received model – it should go align with oil price forecasting.
All coefficients by PVAR estimation are presented in the table (2.4) below:

Table 2.4 Panel vector autoregression estimation results

	Equation (1) $SALES_{it}$ Coef. [Std. Errors]	Equation (2) $OPEX_{it}$ Coef. [Std. Errors]	Equation (3) NOA_{it} Coef. [Std. Errors]
$SALES_{it-1}$	-0.0954 [.5207]	-0.9792*** [0.6028]	0.4094 [0.3299]
$OPEX_{it-1}$	0.2467 [0.3673]	1.0526*** [0 .4272]	-0.5500*** [0.2166]
NOA_{it-1}	1.0540*** [0.5679]	1.2747*** [0.6473]	0.6377** [0.3642]
$Oilprice$	0.0889 [0.4439]	-0.0917 [0.5229]	0.6452*** [0.64523]
$Oilreserves$	1.7775*** [0.7022]	2.0023*** [0.7971]	-0.1898 [0.3806]
$Oilproduction$	0.8613** [0.5788]	0.8497* [0.6105]	-0.1627 [0.4788]
*p<0.1, **p<0.05, ***p<0.01			

As we can see, repercussion of interdependency is highly noticeable in forecasted model, especially for OPEX: its future value is highly dependent on today's SALES, OPEX and NOA performance. Dependency on future prices is necessary only for more precise calculation of NOA. It is even better for analysts, because future oil prices are unknown for now and highly volatile, thereby complicating it's forecasting.

After estimating system of equations (2.19), were received a P-value < 0.0000 for all three equations, which means, we have significant equation and can use them for further analysis. From the equation (1), we can see that only part of coefficients become significant for SALES influence. Both coefficients for SALES and OPEX at time *t-1* is insignificant, therefore, they will not be used for forecasting purposes. Coefficient for NOA at time *t-1* is close to one – 1.054026 with a high level of significance (p<0.01). In contrast with SEM estimation (2.18), oil price is not significant for influence on SALES. This fact was quite surprising. Both oil production and oil reserves showed a significant influence on SALES with coefficients 0.8613 and 1.7775 respectively.

In equation (2) coefficient for SALES is significant (p<0.01) with value of -0.9794. The same situation with significant coefficients for lagged OPEX and lagged NOA. Again, in contrast with SEM estimation (2.2), oil price is not significant for influence on endogenous variable (in our case, OPEX). As in equation (2), both oil production and oil reserves showed a significant influence on SALES with coefficients 0.8497 and 2.0023 respectively.

Last equation (3) is dedicated to finding possible dependency of net operating assets. Actually, by performing estimation, only half of coefficients was found as significant. Coefficient of SALES at time *t* is found as insignificant, the same unexpected result for oil reserves and oil production. Both coefficients for OPEX and NOA at time *t-1* is significant, but with different levels of confidence, $p<0.01$ and $p<0.05$, respectively. As for oil price, finally, at least, for NOA, it become significant with a high level of confidence. We estimated coefficient and decided, which of them we need to consider and forecast for to receive results that are more precise. Nevertheless, for now it is quite uncomfortable to use just number, thereby there was taken a decision to “pack” all receive data into user-friendly interface and to create a computer-forecasting program.

2.2.2 Virtual Basic for Application Program Description

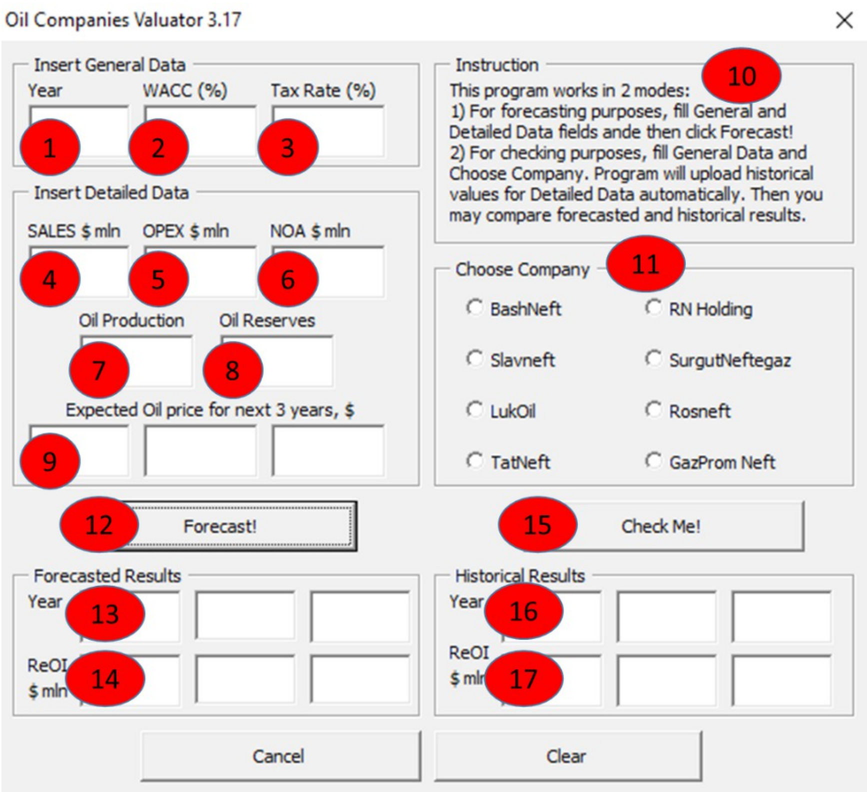
Received results from G2SLS and PVAR estimation are quite difficult for interpretation. Yes, by performing estimation, we know on which variables investor should be focused mostly. Also, we found which variables are not significant, moreover, we built forecasting model with precise structure and dependence on lagged variables. Very important finding is building valuation forecasting align with forecasting of oil prices, because oil price is a very volatile and cannot be plugged into the model.

In a purpose to have stronger practical implementation and for real practical usage of current work, was created a software in a Virtual Basic for Application (VBA) environment. General purpose in in automatization of forecasting process. All we need is to input some company’s variables and run the forecasting process.

The main feature of the current program is possibility to work in two different modes: forecasting and historical checking.

- *Forecasting mode.*
  This mode is created for forecasting purposes of any oil companies. Code is based on PVAR estimation. All received coefficients are plugged in the code. This mode is the best suitable for Russian oil companies, but can work with oil industry in different countries.
- *Historical checking mode.*
  This mode is created for model accuracy testing. Here, forecasting process is using too, but with different data. All we need to do is to choose the one of eight Russian oil companies and year of checking. All historical data automatically will be plugged in the part of detailed companies’ data. Then, forecasting process will be performed with historical data. Finally, there is possibility to check historical residual operating income and forecasted residual operating income based on model, created in current masterwork.

Program interface is presented below:



Picture 2.1 Oil Companies Valuator

1. Here is the part refers to the general data, which should be inputted every time for different valuations. First field devoted to the current year of forecasting as a starting point. It is not crucial part, but for better understanding, current year should be inputted, and forecasting years will be taken as +1, +2, +3 respectively.
2. This box is for weighted average cost of capital. This parameter is different for different companies and even for one company in different periods of time, so it should be inputted manually every time.
3. Last box in general data is for tax rate. Usually it is 20% for Russian companies, but decision of manual input was taken for a case of changing in legislation or working with different countries.
4. This part is devoted for more detailed data about company's performance. This part is filling only for forecasting purposed, not for historical checking of model accuracy. First box is for SALES amount in millions of dollars.
5. Second box in this part is for operational expenses in millions of dollars.
6. Next box represents net operating assets amount expressed in millions of dollars too.
7. Further, here is a place for annual oil production amount in barrels.

8. This box is created for oil reserves amount in barrels. Having many companies with resumption rate of more than 100%, this number is slightly increasing from year to year, but not significantly.
9. Last free fields in this part are devoted to crude oil prices for next three years. As it was mentioned above, this parameter is quite volatile and hard to predict, so oil prices could not be implemented into the model.
10. Here is the short instruction, which explains how to use two different modes of the program.
11. This part is useful for checking mode – it automatically loads and fills empty boxes with companies' parameters. Here is the choice of eight Russian oil companies, whose data was gathered for current research.
12. This button launch forecasting mode.
13. In these three boxes next three forecasting years are shown.
14. Here forecasted residual operating income is shown after forecasting process.
15. Button activates historical checking mode with all historical data loaded for companies' performance and, at the same time forecasting process based on historical data.
16. Here posted historical results and years. To receive any data here, last possible year with available information – 2014.
17. In these three boxes historical residual operating income is shown

Current VBA program was created for more comfortable valuation process. Code in this program is based on PVAR estimation coefficients received during research process. This program is open for further adjustments and code improvement.

Conclusion

Studying of Russian oil industry showed the importance of developing and improvement of existing valuation models. Having almost all existing valuation models universal, that fact makes them inefficient for working with features of different industries. In the first chapter, all main principles of value-based management were studied, as a basement for further research.

While performing current masterwork, main valuation models were studied, divided into groups and their main features are discussed. Then, residual operating income was highlighted from other valuation models and taken for further improvement. Within the model, focus was done on only one variable, which needed to be forecasted – operating income.

For further improvement and possible adjustments, significant part of literature review was devoted to simultaneous equations model. Studying of articles written in this field showed a possibility of implementation of simultaneous equation model into residual operating income model for improvement purposes.

In a developed model three main exogenous variables were taken – sales, operating expenditures and net operating assets. Model was checked on the set of eight biggest Russian oil companies by 3SLS estimation. Estimation results showed an possibility of using developed model. Then, created model was extended by adding exogenous variables – crude oil price, its volatility, oil production, oil reserves and two ratios – EV/EBITDA and EBITDA/Barrel. G2SLS model estimation indicated significance of only oil price, production and reserves coefficients.

Finally, forecasting model was created, based on previously received estimation results. This last step was performed by PVAR estimation tool.

Received forecasting model has almost all necessary information to work with valuation of Russian oil companies. One important note – current mode needed to be used with align of crude oil price forecasting – as it is a very volatile variable, which could not be directly implemented into developed model.

For more strong managerial implication and for more comfortable way of using developed model, there was a decision of creating a VBA program. This tool could be used for valuation purposes, with a set of minimum data which needed to be inputted.

There is a possible ways of improvement of created model and application. However, current masterwork created a basement of using simultaneous equation model within residual operating income model to forecasting the value of Russian oil companies.

References

1. Abrigo, M., Love, I. 2015. *"Estimation of Panel Vector Autoregression in Stata: a Package of Programs*". Discussion Paper. University of Hawai`i at Manoa, 45 p.
2. Angrist, J. D., Imbens, J. W. 1995. *"Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensit".* Journal of the American Statistical Association 90: pp. 434-446.
3. Balestra, P., Varadharajan-Krishnakumar, J. 1987. "*Full information estimations of a system of simultaneous equations with error component structure"*. Econometric Theory 3: pp. 223–246.
4. Baltagi, B. H. 2013. *Econometric Analysis of Panel Data*. Chichester, UK: Wiley, 467 p.
5. Beidas-Strom, S., Pescatori, A. 2014. *"Oil Price Volatility And The Role Of Speculation"*. International Monetary Fund 14: pp. 1-28.
6. Bergmann, I. 2014. *"Accounting based valuation and 'other information' – A simultaneous equations approach"*. Discussion Paper, 2014 International Accounting Section Mid-Year Meeting and Doctoral/New Faculty Consortium, 56 p.
7. Bernard, V. I. 1995. "*The Feltham-Ohlson Framework: Implications for Empiricists"*. Contemporary Accounting Research 11(2): pp. 733–747.
8. Bollen, K. A. 1989. *Structural Equations with Latent Variables*. New York: Wiley, 528 p.
9. Bukhvalov, A. V. 2005. *"Corporate governance as an object for scientific researches"* ["Korporativnoe upravlenie kak ob"ekt nauchnyh issledovanij"]. Russian Management Journal 3(3), pp 81-96.
10. Callen, J. L., Segal, D. 2005. "*Empirical Tests of the Feltham-Ohlson (1995) Model"*. Review of Accounting Studies 10(4): pp. 409–429.
11. Copeland, T. E., Koller, T., Murrin, J. 1995. *Valuation: measuring and managing the value of companies*. New York: Wiley, 558 p.
12. Damodaran, A. 2002. *Investment Valuation*. New York: Wiley, 921 p.
13. Davidson, R. Mackinnon, J. D. 1993. *"Estimation and inference in econometrics"*. New York: Oxford University. 871 p.
14. Destri, M. L., Mina, A., Picone, P. M. 2012 *Corporate Social Performance: Paradoxes, Pitfalls and Pathways To The Better World*, Charlotte, NC, USA: Information Age Publishing, 657 p.
15. Devold, H. 2013. *Oil and gas production handbook: an introduction to oil and gas production, transport, refining and petrochemical industry*. Oslo: ABB, 921 p.
16. Elliot, J. W., Uphoff, H. L. 1972. *"Predicting the near term profit and loss statement with an econometric model: a feasibility study"*. Journal of Accounting Research 13: pp 1499-1526.
17. Fama, E. F., French. K. R. 2004. "*The Capital Asset Pricing Model: Theory and Evidence"*. Journal of Economic Perspectives 18(3): pp 25–46.
18. Greene, W. H. 2012. *Econometric Analysis*. New York: Pearson, 1241 p.

19. Gurvich, E. T. 2004. "Macroeconomic valuation of the role of Russian oil industry" ["Makroehkonomicheskaya ocenka roli rossijskogo neftegazovogo sektora"]. Questions of economics 10: pp. 1-36.
20. Holtz-Eakin, D., Newey W., Rosen H. S 1988. "*Estimating vector autoregressions with panel data"*. Econometrica, 56(6): pp. 1371-1395.
21. Hausman, J. A. (November 1978). *"Specification Tests in Econometrics"*. Econometrica 46(6): pp. 1251–1271.
22. Jensen M. 2001. "*Value Maximisation, Stakeholder Theory, and the Corporate Objective Function"*. European Financial Management 7(3): pp. 297-317.
23. Koller, T., Goedhart, M., Wessels, D. 2010. *Valuation: Measuring and Managing the Value of Companies*. New Jersey: John Wiley & Sons, 840 p.
24. Kudrin, A. 2014. *"The Influence of Oil and Gas Revenues on Russia's Monetary and Credit Policy"*. Working Paper, Gaidar Institute for Economic Policy, 27 p.
25. Lambert M., Hübner G. 2014 "*Size matters, book-‑to-market does not! The F&F empirical CAPM revisited"*. 12th Paris Finance Meeting, 23 p.
26. Love, I., Zicchino, L. 2006. "*Financial development and dynamic investment behavior: Evidence from panel VAR"*. The Quarterly Review of Economics and Finance, 46(2): pp. 190-210.
27. McFadden, D. 1999. *Chapter 6. Simultaneous equations*. Resource document. Lecture notes for Economics, Berkeley, 39 p.
28. Nissim, D., Penman, S. 2001. *"Ratio Analysis and Equity Valuation: From Research to Practice"*. Review of Accounting Studies 6 (1): pp. 109-154.
29. Ohlson, J.A. 1995. "*Earnings, Book Values, and Dividends in Equity Valuation"*. Contemporary Accounting Research 11(2): pp. 661-687.
30. Ohlson, J.A. 2001. "*Earnings, Book Values, and Dividends in Equity Valuation: An Empirical Perspective".* Contemporary Accounting Research 18(1): pp. 107-120.
31. Osmundsen, P., Asche, P., Misund, B., Mohn, K. 2006. "*Valuation of International Oil Companies"*. The Energy Journal 27(3): pp. 49-64.
32. Penman, S. H. 2010. *Financial Statement Analysis and Security Valuation*. New York: McGraw-Hill/Irwin, 768 p.
33. Penman, S.H., Sougiannis, T. 1998. "*A Comparison of Dividend, Cash Flow, and Earnings Approaches to Equity Valuation".* Contemporary Accounting Research 15(3): pp. 343- 383.
34. Reiss, P. C., Wolak, F. A. 2007. *"Structural econometric modeling: rationales and examples from industrial organization"*. Handbook of Econometrics 6: pp. 4280-4317
35. Richardson, S., Tuna, I., Wysocki, P. 2010. *"Accounting Anomalies and Fundamental Analysis: A Review of Recent Research Advances"*. Working Paper. New York Conference, 37 p.
36. Rutgauzer, V. M. 2007. *Business Valuation* [Ocenka stoimosti biznesa]. Moscow: Maroseyka, 448 p.
37. Sims, C.A. 1980. "*Macroeconomics and realit*y". Econometrica, 48(1), pp. 1-48.

38. Soliman, M. T. 2008. *"The Use of DuPont Analysis by Market Participants"*. The Accounting Review 83 (3): pp. 823-853.
39. Volkov, D. L. 2008. *"Theory of value-based management: financial and accounting aspects"* ["Teoriya cennostno-orientirovannogo menedzhmenta: finansovyj i buhgalterskij aspekty"]. Saint Petersburg: GSOM, 320 p.
40. Volkov, D. L., Rudenko, A. 2009. *"Value-based approach in managing of international high-tech companies*" Corporate Finance 2(10): pp. 5-27.
41. Wallace, J. S. 2003. *"Value maximization and stakeholder theory: compatible or not?"* Journal of Applied Corporate Finance 15(3): pp. 120–127.
42. Wright, C. J., Gallun, R. A. 2008. *Fundamentals of Oil & Gas Accounting.* New York: PenWell Books, 750 p.