Vitaly Buylov

Practical Application of Econometric Modelling to Market Approach Method Using Tesla Motors, Inc. as an Example

Metropolia University of Applied Sciences

Bachelor of Business Administration

European Business Administration

Thesis

April 2022

Abstract

Author(s):	Vitaly Buylov
Title:	Practical application of econometric modeling to market approach method using Tesla Motors, Inc. as an example.
Number of Pages:	60 pages
Date:	27 April 2022
Degree:	Bachelor of Business Administration
Degree Programme:	European Business Administration
Specialization option:	Quantitative Business Analytics
Instructor(s):	Kevin McIntire, Senior Lecturer Michael Keaney, Senior Lecturer

The aim of this study is to some extent a reality check for one of the most commonly used contemporary equity valuation methods. Particular interest is the application of the multiple linear regression method to the market valuation approach. Tesla Motors Inc. is used as an example of a company under valuation.

The paper provides a brief overview of the market valuation approach as well as a detailed overview of the process of multilinear regression modelling including but not limited to the statistical inference after logarithmic transformation.

The rationale for choosing Tesla Motors Inc. lies behind its uniqueness which together with the combined econometric market valuation approach is expected to drive this method to its limitations.

As a result, the study shows that although this combined approach under particular conditions can yield some trustworthy results valuer should not entirely rely on that method only. Instead, it is recommended to use it in a combination with other valuation methods such as discounted cash flow approach.

Keywords: Contemporary Corporate Equity Valuation Issues

Contents

Glossary

BV	Book value.
EBITDA	Earnings before interest taxes depreciation and amortization.
EPS	Earnings per share ratio.
EV	Electric vehicle.
GNU	General public license.
GRETL	Gnu regression, econometrics, and time-series library.
INSC	International valuation standards council.
MLRM	Multiple linear regression model.
OLS	Ordinary least squares.
P/B	Price to book ratio.
P/E	Price to earnings ratio.
P/S	Price to sales ratio.
TTM	Trailing twelve months.
VIF	Variance inflation factor.

1 Introduction

On Monday, October 25, 2021, after the news of a deal with international rental car company Hertz Global Holdings Inc (HTZ), the market capitalization of Tesla Inc (TSLA) reached a valuation of USD 1 trillion first time in the whole history of the automotive industry.

“The rally in Tesla’s shares has lifted the overall stock market value of Elon Musk’s electric carmaker to over USD 1.1 trillion, making it one of the most valuable companies in the world. This year alone it has added almost USD 475 billion in market capitalization, equal to a Procter & Gamble, a JPMorgan — or two McDonald’s.” (Wigglesworth, 2021)

The same concerns about Tesla valuation were addressed by Herath:

“It is fair to say that Tesla is one of the most polarizing stocks of our time. On TipRanks, the 1-year prediction of the stock price ranges from a low of USD 67 to a high of USD 1,200. That is a staggering difference of about 20x. Even when looking for guidance from the most esteemed analysts of our time, we get extremely polarized views.” (Herath, 2021)

Financial advisers, equity portfolio managers, and chartered financial analysts are head over heels about Tesla’s perspective growth. Lee uses the word “growth” 6 times once in each sentence and goes on and on claiming that “Tesla as an early-stage growth stock. The company … could be positioned for growth well into the future. Tesla’s growth potential is consistent with our growth investment process, which aims to find companies early in their life cycles. We think investing in early-stage companies provides an opportunity for investors to capture higher growth—and sustain it longer. …Identifying above-average growers early in their lifecycles provides an opportunity to compound that growth overtime.” (Lee, 2020)

At the same time, Tesla’s global market share started to go down as competition heats up:

“A recent report from Canalys states that Tesla’s long-term market share leadership is under threat. Per its data, Tesla’s global EV market share for the first six months of 2021 was 15%. This compares unfavourably with 19% share in the first half of 2020.” (Singhi, 2021)

1.1 Motivation.

The purpose of this study is to determine if contemporary theory on business valuation methods, particularly the theory on market approach method, can justify current Tesla’s stock price. In other words, we are trying to answer the question of how much (if so) Tesla Inc. is over-/ or under-priced according to the applied market approach method in valuation.

“The current mania for electric-vehicle makers, from Tesla to Lucid to Rivian, is reminiscent of the dot-com boom and bust.” (Cassidy, 2021)

“Returning to the lessons of 2000, great companies can be terrible investments. The unfortunate object lesson came from Cisco Systems Inc., briefly the world’s largest company, and in 2000 the dominant provider of the internet’s physical infrastructure. This provides an easy bear case for Tesla. Yes, it’s a great company, and yes electric vehicles are the future. But you still don’t want to buy at this price. Just look at Cisco.” (Authers, 2021)

To avoid the next dot-com crisis, private investors, business analysts, and financial institutions alike should be more than ever sceptical and prudent when considering investing in companies like Tesla.

“Extensive disputes around valuation outcomes are detrimental as ultimately the valuation object’s value can be affected by the time and attention a dispute demands. Additionally, strong deviations in assumptions may result in an under- or overvaluation of a company, especially in times of high uncertainty, which influences the quality and soundness of investment decisions (i.e., buy or sell), as well as contribute to potential capital destruction.“ (Broekema, Marc J. R. et al., 2020)

Even though the topic of the research is related to the hyper-popular Tesla Motors Inc. and therefore probably the most common (preferred) thesis choice among students pursuing the career path in the quantitative field of business, this study anyway is directed to bring some unique perspective on the problem of business valuation in these challenging times of turbulent markets and hopefully will contribute to applied business valuation practice.

1.2 Theoretical framework.

The conceptual basis of this thesis revolves around the literature on two general theoretical aspects. Namely, they are contemporary corporate valuation methods of particular interest of which are the market (relative) approach on one hand and advanced econometric theory on multiple linear regression on the other hand. Both of them alike are concisely covered in subsequent chapters of this very research and combined they represent the essence of the methodology applied in this work.

1.3 Research methodology and thesis structure.

The study is based on applied quantitative method which we will use to explore, present and describe the results. For that, we will use the applied OLS method in multiple linear regression modelling. The rationale behind using this particular statistical approach is covered below.

2 Value. Valuation framework. Behavioural perspective.

As Bernstrom claims, “a common misconception is that, in the world of business valuation, only one single universal value exists. Unfortunately, that is not the case. There is a whole variety of different types of values available as well as definitions of value. For this reason, before even working on any spreadsheets, it is very important to clearly define what value we are looking for and why.” (Seth Bernstrom, 2014)

Therefore, we first need to tackle the question of the valuation framework. There are many valuation frameworks in the world today, one of which is provided by the International Valuation Standards Council (IVSC), which is a non-profit organization that provides standards to be used by the valuation profession.

The validity of these standards depends on the extent to which they are recognized and applied. IVSC is widely recognized by educational institutions, among which are: Al Muheet Institute, Ankara University, Department of Real Estate Development and Management, ICMAI Registered Valuers Organisation, Indian Institute of Corporate Affairs, Institute for Mergers, Acquisitions and Alliances GmbH, Institute of Finance, School of Business, University of Applied Sciences & Arts, NW Switzerland, International Institute of Business Valuers, Italian Association of Professors in Accounting & Business Administration (SIDREA), Leventhal School of Accounting, University of Southern California, to name just a few.

“Members of the IVSC represent leading organisations in the mission to develop consistent, quality valuation standards, and a global valuation profession. Membership is open to valuation end users, service providers, professional and accrediting bodies, educators, and regulators.” (IVSC, 2022)

IVSC gives its definition of value which is “the opinion resulting from a valuation process that is compliant with IVS. It is an estimate of either the most probable monetary consideration for an interest in an asset or the economic benefits of holding an interest in an asset on a stated basis of value. “ (IVSC, 2022)

- Valuation

“Valuation - the act or process of determining an opinion or conclusion of value of an asset on a stated basis of value at a specified date in compliance with IVSC.” (IVSC 2022)

- Valuation Approach.

Valuation approach – “in general, a way of estimating value that employs one or more specific valuation methods”. (IVSC 2022)

According to the framework, “the principal valuation approaches are:

(a) market approach,
(b) income approach, and
(c) cost approach.” (IVSC 2022)

In this study market approach will be tested for its applicability and efficiency when valuing such companies as Tesla Inc.

- Valuation Method.

“Valuation method - within valuation approaches, a specific way to estimate a value.” (IVSC, 2022)

In this work market approach will be combined with multiple linear regression modelling using OLS estimators.

- Purpose of the valuation.

“Purpose of the valuation - the purpose for which the valuation assignment is being prepared must be identified as it is important that valuation advice is not used out of context or for purposes for which it is not intended.” (IVSC, 2022)

In the case of this study, the purpose of the valuation is strictly academic. What we intend to do here is to create a common econometric model for Tesla’s peer group using market capitalization as an explained variable and different metrics and ratios given by market valuation as explanatory variables and then compare the model's valuation for Tesla with its current market capitalization. Thus, we set ourselves the goal of comparing theory on market approach for corporate valuation within IVSC framework with the real-world today’s equity market value of Tesla (actualised on 30.01.2022)

- Valuation biases. A behavioural perspective.

As Damodaran claims “in theory, we start with the financial fundamentals and move “objectively” from the numbers to the value of the firm, making reasonable assumptions along the way. In practice, though, valuations are not just subjective but are contaminated by biases that analysts bring to the process. … In fact, it is not uncommon to see analysts change their assumptions to move their valuations closer to the stock price. … There is evidence that when data are presented sequentially, the most recent data are weighted too much (relative to its importance) and less recent data too little. … As a consequence, we tend to overvalue companies after good years and undervalue companies after bad years. … There is some evidence that analysts who form a perception of what the fair value is early in the process then tend to model the data to confirm that perception. As a result of these biases, we would argue that in many valuations, the value gets set first and the valuation follows." (Damodaran, 2014)

- Valuing Growth Companies. A behavioural Perspective.

In addition, Damodaran warns that "in theory, we should expect to see larger valuation errors with growth companies than with mature companies, because there is more firm-specific uncertainty that we face in valuing growth companies, insofar as we have to estimate how long growth will last and how high growth will be during the period. In practice, we generally find support for this hypothesis but we also find that there is more bias in the valuation of growth companies. In particular, there is evidence to suggest that high growth (and high PE) stocks tend to earn returns that are too low and are thus priced too high, relative to low-growth stocks." (Damodaran, 2014)

3 Market valuation approach.

3.1 Definition.

"The market approach provides an indication of value by comparing the asset with identical or comparable (that is similar) assets for which price information is available." (IVSC, 2022)

Similarly, Bernstrom defines this method as follows: "the market approach aims to derive the value of a company based on how similar firms are priced on the stock exchange or through company transactions. Consequently, the pricing of the valuation subject will implicitly be dependent upon other actors' assessment of future growth potential, profitability, risk profile (cost of capital), etc. for the valuation subject in question, which may or may not be appropriate. Differences between the comparator group of companies and the valuation subject at issue as regards the size and nature of their operations, among other things, will justify correspondingly different levels of business risk, growth potential, margins, etc. These differences must therefore be considered when justifying different levels of value, i.e. when justifying the relevant or appropriate value multiple to be applied to the subject company." (Seth Bernstrom, 2014)

Finally, as Damodaran states: "in relative valuation, the objective is to value assets, based on how similar assets are currently priced in the market." (Damodaran, 2014)

3.2 Method procedure.

According to Fazzini, "the application of the market multiple method involves the following steps:

a) selection of a peer group;
b) choice of multiple;
c) application of multiple to the target company." (Fazzini, 2018)

3.3 The peer group.

Bernstrom claims that "in most cases, it is natural to seek peers within the same geographical area and the same sector or industry as the valuation subject. If, for example, we are valuing a European car manufacturer, it would be natural to seek peers within that current pool of companies." (Seth Bernstrom, 2014)

Fazzini says that "in the market multiple method (market approach), the selection of a homogeneous peer group is essential to obtain a reliable value of the target company…. the choice of the sample requires analysing both quantitative and qualitative variables." (Fazzini, 2018)

Therefore, let’s start building our reference pool of companies from globally traded public automakers with a market capitalization of more than USA 20 billion (Figure 1). For that reason, data from Yahoo! Finance will be retrieved.

“Yahoo! Finance is a media property that is part of the Yahoo! network. It provides financial news, data, and commentary including stock quotes, press releases, financial reports, and original content.” (Yahoo! Finance, 2022)

Symbol	Company Name
TSLA	Tesla, Inc.
RACE	Ferrari N.V.
TM	Toyota Motor Corporation
RIVN	Rivian Automotive, Inc.
GM	General Motors Company
TTM	Tata Motors Limited
XPEV	XPeng Inc.
HMC	Honda Motor Co., Ltd.
LI	Li Auto Inc.
LCID	Lucid Group, Inc.
NIO	NIO Inc.
F	Ford Motor Company
STLA	Stellantis N.V.

Figure 1. Top companies from automakers. (Yahoo! Finance, 2022)

It is not a secret that Tesla Inc. is one of the most unique companies of our time, one of a kind and therefore it is challenging to find another similar company not to mention a statistically appropriate sample of companies for sufficient data analysis. However, there are workarounds that we can make to overcome the problem of sample limitation by including closely relative companies.

The Indo-Asian News Service (IANS) claimed that according to Tesla CEO Elon Musk "Tesla is not just a car maker but an AI robotics firm" (IANS, 2021). Therefore we can expand our peer group by including additional 25 companies from the technology sector which worth more than USA 20 billion. (Figure 2)

Subramanian mentioned that Charles Janac a CEO of chip tech company Arteris shared that in his opinion “Tesla is not necessarily a car company — It's an internet of cars company. They control their software architecture very well, and ... they make some of their own chips." (Subramanian, 2021) Therefore, we can include in the pool another 25 companies from the industrial sector which also have the USA 20+ billion worth. (Figure 3).

According to Dunn, “Tesla invested roughly 18% of its revenue into research and development last year”. (Dunn, 2016) That’s nearly three times as high as most traditional car companies.” Such big spendings on R&D is in line with the spendings of the healthcare sector many of which “spend as much as 25% of their revenue on R&D” (Investopedia, 2022). So that we can add the last 25 companies from healthcare (Figure 4) and thereby combine all the companies into one sample of 88 companies.

- The choice of multiples and their econometric application.

Since this study uses an econometric method, particularly multilinear regression modelling there is a special approach needed for the selection of explanatory variables (multiples). This approach aims to mitigate the problem of model misspecification therefore it has been moved to section 5 (secondary research data collection).

Symbol	Company Name
AAPL	Apple Inc.
MSFT	Microsoft Corporation
NVDA	NVIDIA Corporation
TSM	Taiwan Semiconductor Manufacturing Company Limited
ASML	ASML Holding N.V.
AVGO	Broadcom Inc.
CSCO	Cisco Systems, Inc.
ORCL	Oracle Corporation
ADBE	Adobe Inc.
ACN	Accenture plc
CRM	salesforce.com, inc.
INTC	Intel Corporation
AMD	Advanced Micro Devices, Inc.
TXN	Texas Instruments Incorporated
QCOM	QUALCOMM Incorporated
INTU	Intuit Inc.
SAP	SAP SE
SONY	Sony Group Corporation
IBM	International Business Machines Corporation
AMAT	Applied Materials, Inc.
NOW	ServiceNow, Inc.
INFY	Infosys Limited
SHOP	Shopify Inc.
MU	Micron Technology, Inc.
TEAM	Atlassian Corporation Plc

Figure 2. Top 25 companies from the technological sector. (Yahoo! Finance, 2022)

Symbol	Company Name
UPS	United Parcel Service, Inc.
UNP	Union Pacific Corporation
RTX	Raytheon Technologies Corporation
HON	Honeywell International Inc.
DE	Deere & Company
LMT	Lockheed Martin Corporation
CAT	Caterpillar Inc.
BA	The Boeing Company
GE	General Electric Company
ADP	Automatic Data Processing, Inc.
CNI	Canadian National Railway Company
MMM	3M Company
CSX	CSX Corporation
CP	Canadian Pacific Railway Limited
NOC	Northrop Grumman Corporation
NSC	Norfolk Southern Corporation
GD	General Dynamics Corporation
WM	Waste Management, Inc.
ITW	Illinois Tool Works Inc.
ABB	ABB Ltd
ETN	Eaton Corporation plc
EMR	Emerson Electric Co.
FDX	FedEx Corporation
TRI	Thomson Reuters Corporation
JCI	Johnson Controls International plc

Figure 3. Top 25 companies from the industrial sector. (Yahoo! Finance, 2022)

Symbol	Company Name
UNH	UnitedHealth Group Incorporated
JNJ	Johnson & Johnson
PFE	Pfizer Inc.
ABBV	AbbVie Inc.
LLY	Eli Lilly and Company
NVO	Novo Nordisk A/S
TMO	Thermo Fisher Scientific Inc.
MRK	Merck & Co., Inc.
ABT	Abbott Laboratories
DHR	Danaher Corporation
AZN	AstraZeneca PLC
NVS	Novartis AG
BMY	Bristol-Myers Squibb Company
MDT	Medtronic plc
AMGN	Amgen Inc.
CVS	CVS Health Corporation
SNY	Sanofi
ANTM	Anthem, Inc.
GSK	GlaxoSmithKline plc
ISRG	Intuitive Surgical, Inc.
SYK	Stryker Corporation
ZTS	Zoetis Inc.
CI	Cigna Corporation
GILD	Gilead Sciences, Inc.
BDX	Becton, Dickinson and Company

Figure 4. Top 25 companies from the healthcare sector. (Yahoo! Finance, 2022)

3.4 Method's drawbacks

Here we should also mention that although this method is widely common among professional valuers and analysts it, nevertheless, has some negative attributes which were addressed by Hooke. One of the drawbacks of this approach is that "many subject businesses lack a set of true comparable firms, so there is little to which to relate them." (Jeffrey C. Hooke, 2010). This claim is even more apparent to our research since the company we are analysing is truly a pioneer in the EV industry and as such doesn't have close relatives to compare with.

It is hard to disagree with Hooke over another negative attribute of this method related to market inefficiency and as a result the lack of a "yardstick to indicate whether the entire group of comparable is properly valued." There he also added that "during the dot-com boom, the pricing of the entire Internet sector was inflated." (Jeffrey C. Hooke, 2010).

Damodaran applies the same reasoning while choosing between types of models. "When using relative valuation, it is dangerous to base valuations on multiples where the differences across firms cannot be explained well using financial fundamentals—growth, risk, and cash flow patterns. One of the advantages of using the regression approach … is that the R^2 and t-statistics from the regressions yield a tangible estimate of the strength (or weakness) of this relationship." (Damodaran, 2014)

4 OLS method. Simple and multiple linear regressions. Model specification. Inference.

4.1 Definitions

“The simplest linear regression technique. **OLS** involves fitting a linear equation with the coefficients chosen to minimize the sum of squares of residual errors.” (John Black, et al., 2009)

Residual is a “deviation between the data and the fit, it is also a measure of the variability in the response variable not explained by the regression model.” (Douglas C. Montgomery, et al., 2012)

Ceteris paribus, literally "holding other things constant," is a Latin phrase that is commonly translated into English as "all else being equal." (Liberto, 2021)

Montgomery defines a “**simple linear regression** model, that is, a model with a single regressor x that has a relationship with a response y that is a straight line. This simple linear regression model is:

$$y = \beta_0 + \beta_1 * x + \varepsilon;$$

where the intercept β_0 and the slope β_1 are unknown constants and ε is a random error component.” (Douglas C. Montgomery, et al., 2012)

Adkins compares simple and **multiple linear regression** and defines the latter as “an extension of the simple model. The main difference is that the multiple linear regression model contains more than one explanatory variable. This changes the interpretation of the coefficients slightly and requires another assumption. The general form of the model is shown in the equation below:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdot\cdot\cdot + \beta_K x_{iK} + e_i; \quad i = 1, 2, . . . , N$$

where y_i is your dependent variable, x_{ik} is the i^{th} observation on the k^{th} independent variable, $k = 1, 2, . . . , K$, e_i is a random error, and $\beta_1, \beta_2, . . . , \beta_K$ are the parameters you want to estimate." (Lee C. Adkins, 2011)

"$\mathbf{R^2}$ is the proportion of the sample variation in the dependent variable explained by the independent variables, and it serves as a **goodness-of-fit** measure. It is important not to put too much weight on the value of R^2 when evaluating econometric models." (Wooldridge, 2015)

4.2 Applied econometric analysis structure.

In Figure 5. there is a schematic description of the procedure of econometric analysis.

According to the scheme pictured below, we are going to test the theory on the precision of market (relative) valuation method by building econometric models using free source cross sectional data from our secondary research followed by making estimations and using the model for computing the market capitalization of Tesla Motors Inc.

If specification test shows that the model is adequate, we can to some extent and with some limitations draw a conclusion about the efficiency of market valuation method applied to such companies like Tesla Motors Inc. It is the essence of this study.

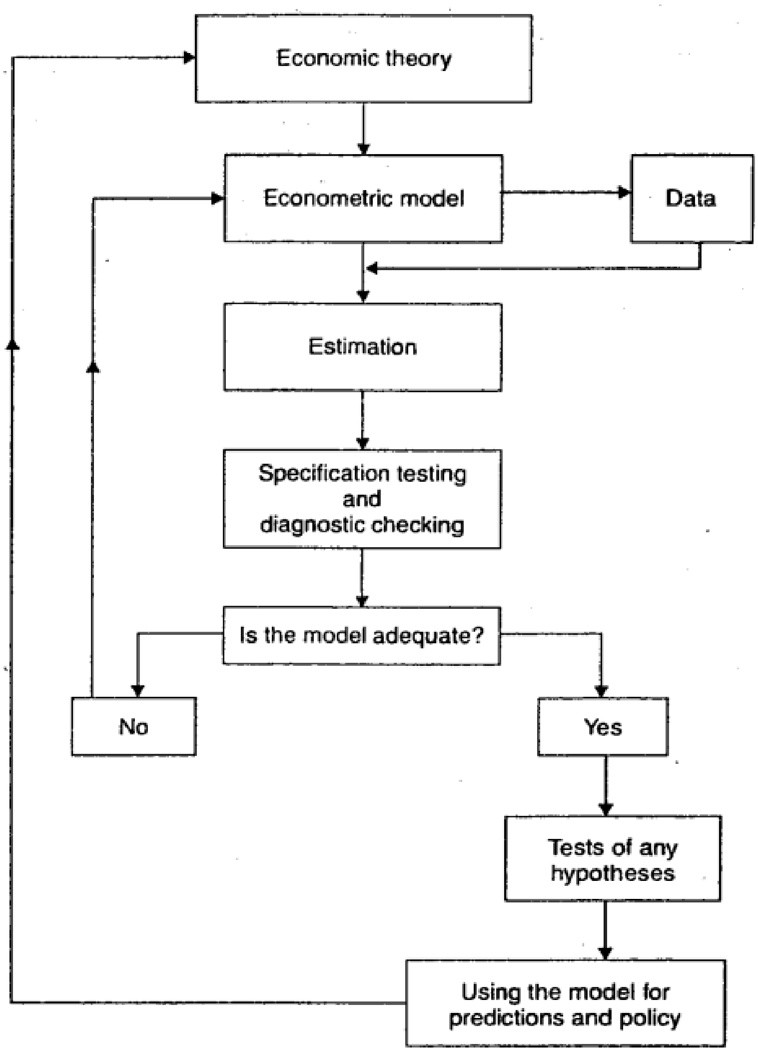

Figure 5. The stages of applied econometric analysis. (Dimitrios Asteriou & Stephen G. Hall, 2007)

4.3 Econometric method selection.

We opted out to use the ordinary least squares (OLS) method in multiple linear regression modelling as a tool for econometric analysis.

The rationale behind using multiple linear regression instead of simple linear regression is that, although the effect of each independent variable (such as revenue, growth rates, EBITDA, BV etc.) could be estimated by a simple linear regression of Market capitalizations on each explanatory variable separately these results may be misleading because the explanatory variables can be related. For instance, the revenue effect on market capitalization is partly caused by the revenue effect on EBITDA. This mutual dependence is taken into account by a multiple regression model that includes more than one independent variable.

Wooldridge agrees with above mentioned claim that “multiple regression analysis is more amenable to ceteris paribus analysis because it allows us to explicitly control for many other factors that simultaneously affect the dependent variable. Multiple regression models can accommodate many explanatory variables that may be correlated, we can hope to infer causality in cases where simple regression analysis would be misleading. If we add more factors to our model that are useful for explaining “y”, then more of the variation in “y” can be explained. Thus, multiple regression analysis can be used to build better models for predicting the dependent variable.” (Wooldridge, 2015)

Wooldridge also states that “The multiple regression model is still the most widely used vehicle for empirical analysis in economics and other social sciences. Likewise, the method of ordinary least squares is popularly used for estimating the parameters of the multiple regression model.” (Wooldridge, 2015)

4.4 Model building techniques.

According to Fabozzi “there are three methods that are commonly used to determine the suitable independent variables to be included in a final regression model.”

They are:

1. Stepwise inclusion regression method;
2. Stepwise exclusion regression method;

3. Standard stepwise regression method.

“In the stepwise inclusion regression method, we begin by selecting a single independent variable. It should be the one most highly correlated (positive or negative) with the dependent variable. After inclusion of this independent variable, we perform an F-test to determine whether this independent variable is significant for the regression.” (Fabozzi, et al., 2014)

“The stepwise exclusion regression method mechanically is basically the opposite of the stepwise inclusion method. That is one includes all independent variables at the beginning. One after another of the insignificant variables are eliminated until all insignificant independent variables have been removed. The result constitutes the final regression model.” (Fabozzi, et al., 2014)

“The standard stepwise regression method involves introducing independent variables based on significance and explanatory power and possibly eliminating some that have been included at previous steps. The reason for elimination of any such independent variables is that they have now become insignificant after the new independent variables have entered the model.” (Fabozzi, et al., 2014)

4.5 Model adequacy checking. Assumptions of MLRM and their consequences if violated.

- Assumptions of MLRM.

According to Montgomery “the major assumptions for multiple linear regression analysis are as follows:

1. The relationship between the response y and the regressors is linear, at least approximately.
2. The error term ε has zero mean.
3. The error term ε has constant variance σ^2.
4. The errors are uncorrelated.
5. The errors are normally distributed." (Douglas C. Montgomery, et al., 2012)

Montgomery also warns that "we should always consider the validity of these assumptions to be doubtful and conduct analyses to examine the adequacy of the model we have tentatively entertained. … gross violations of the assumptions may yield an unstable model in the sense that a different sample could lead to a totally different model with opposite conclusions. We usually cannot detect departures from the underlying assumptions by examination of the standard summary statistics, such as the t or F statistics, or R^2 . These are "global " model properties, and as such they do not ensure model adequacy." (Douglas C. Montgomery, et al., 2012)

The last number 6 assumption "is the structure or interaction of the independent variables. The statistical term used for the problem that arises from the high correlations among the independent variables used in a multiple regression model is multicollinearity or, simply, collinearity. Tests for the presence of multicollinearity must be performed after the model's significance has been determined and all significant independent variables to be used in the final regression have been determined." (Fabozzi, et al., 2014)

- Consequences of the violation of MLRM assumptions.

Violation of 1st of the above-mentioned assumptions of linearity leads to "violation of assumption one creates problems which are in general called misspecification errors, such as wrong regressors, nonlinearities and changing parameters." (Dimitrios Asteriou & Stephen G. Hall, 2007)

Violation of 2^{nd}, 3^{rd} and the 5^{th} of the above-mentioned assumptions “leads to the inferential statistics of a regression model (i.e .. t-stats, F-stats, etc.) not being valid. Therefore, it is quite essential to test for normality of residuals.” (Dimitrios Asteriou & Stephen G. Hall, 2007)

Violation of 4^{th} of the above-mentioned assumptions leads to autocorrelation under which “the estimated variance of the regression coefficients will be biased and inconsistent and will be greater than the variances of estimate calculated by other methods, therefore, hypothesis testing is no longer valid. In most of the cases, R^2 will be overestimated (indicating a better fit than the one that truly exists). The t- and F-statistics will tend to be higher.” (Ullah, 2020)

Finally, violation of 6^{th} of the above-mentioned assumptions of non-multicollinearity is “a serious problem that may dramatically impact the usefulness of a regression model” which is “near - linear dependence among the regression variables” (Douglas C. Montgomery, et al., 2012). This “results in large variances and covariances for the least - squares estimators of the regression coefficients. This implies that different samples taken at the same could lead to widely different estimates of the model parameters.” (Douglas C. Montgomery, et al., 2012)

To check if these assumptions hold we need to conduct the following tests.

4.5.1 Linearity test.

There are two types of tests on linearity graphical and analytical (quantitative).

- Graphical approach.

As stated by Fabozzi “to test for linearity, a common approach is to plot the regression residuals on the vertical axis and values of the independent variable on the horizontal axis. This graphical analysis is performed for each independent variable. What we are looking for is a random scattering of the residuals around zero. If this should be the case, the model assumption with respect to the residuals is correct. If not, however, there seems to be some systematic behaviour in the residuals that depends on the values of the independent variables. The explanation is that the relationship between the independent and dependent variables is not linear.” (Fabozzi, et al., 2014)

- Formal test. RESET test.

According to Asteriou “one of the most commonly used tests for general misspecification is Ramsey's (1969) Regressions Specification Error Test (RESET) as with many tests this has both an F-form and an LM form. … The RESET test involves including various powers of Y as proxies for X_2^2 that can capture possible non-linear relationships. Before implementing the test we need to decide how many terms we will include in the expanded regression. There is no formal answer to this selestion, but in general the squared and cubed terms have proven to be useful in most applications. … Then the situation boils down to a regular F-type test for the additional explanatory variables Y^2 and Y^3. If one or more of the coefficients are significant then this is evidence of general misspecification”. (Dimitrios Asteriou & Stephen G. Hall, 2007)

This particularly can be used as evidence of non-linear relationships between explained variables and regressors.

Gretl software provides its users a possibility to conduct RESET test, in both the squared and cubed terms.

4.5.2 Heteroscedasticity test.

“There are many tests of the null hypothesis of homoskedasticity that have been proposed elsewhere. Two of these, based on Lagrange multipliers, are particularly simple to do and useful. The first is sometimes referred to as the Breusch-Pagan (BP) test. The second test is credited to White.” (Lee C. Adkins, 2011)

In this study Breusch-Pagan test will be used. “The null and alternative hypotheses for the Breusch-Pagan test are:

$$H_0 : \sigma_i^2 = \sigma^2$$

$$H1 : \sigma_i^2 = h(\alpha_1 + \alpha_2 z_i^2 + \ldots \alpha_s z_{iS})$$

The null hypothesis is that the data are homoskedastic. The alternative is that the data are heteroskedastic in a way that depends upon the variables z_i^s, i = 2, 3, . . . , S. These variables are exogenous and correlated with the model’s variances. The function h(), is not specified. It could be anything that depends on its argument, i.e., the linear function of the variables in z.” (Lee C. Adkins, 2011)

Gretl software package can conduct the Breusch-Pagan test.

4.5.3 Normality test.

Adkins suggests the “Jarque-Bera test for normality which is computed using the skewness and kurtosis of the least squares residuals. To compute the Jarque-Bera statistic, you’ll first need to estimate your model using least squares and then save the residuals to the data set.” (Lee C. Adkins, 2011)

“If the computed p-value of the JB statistic in an application is sufficiently low, which will happen if the value of the statistic is very different from 0, one can reject the hypothesis that the residuals are normally distributed. But if the p-value is reasonably high, which will happen if the value of the statistic is close to zero, we do not reject the normality assumption.” (Gujarati, 2004)

Gretl software package includes the Jarque-Bera test on normality.

4.5.4 Multicollinearity analysis.

To test if there is a correlation between regression variables statistics for individual coefficients should be analysed. “The most common of these is the variance inflation factor (VIF)… Setting a cutoff value for VIF which we conclude multicollinearity is a “problem” is arbitrary and not especially helpful. Sometimes the value 10 is chosen: if VIF is above 10, then we conclude that multicollinearity is a problem.” (Wooldridge, 2015)

Gretl software includes the possibility to easily conduct multicollinearity test.

4.6 Tests on joint and individual significance of the model regressors.

- Joint significance test.

As suggested by Fabozzi “to test whether the entire model is significant, we consider two alternative hypotheses. The first, our null hypothesis H_0, states that all regression coefficients are equal to zero, which means that none of the independent variables play any role. The alternative hypothesis H_1, states that at least one coefficient is different from zero. More formally,

$$H_0: \beta_0 = \beta_1 = \ldots = \beta_k = 0;$$
$$H_1: \beta_j \neq 0 \text{ for at least one } j \in \{1,2,\ldots,k\}$$

In the case of a true null hypothesis, the linear model with the independent variables we have chosen does not describe the behaviour of the dependent variable.“ (Fabozzi, et al., 2014).

- Individual significance test.

“Suppose we have found that the model is significant. Now, we turn to the test of significance for individual independent variables. Formally, for each of the k independent variables, we test

$$H_0: \beta_j = 0; \quad H_1: \beta_j \neq 0;$$

conditional on the other independent variables already included in the regression model. The appropriate test would be the t-test.” (Fabozzi, et al., 2014)

By reading p-values for consequent tests we can decide to reject or not to reject null hypothesis.

4.7 Goodness of fit. Adjusted R^2. Sample size and limits on the number of predictors.

- Adjusted Goodness of fit. Adjusted R^2

Since the type of model used in this study is multiply linear the adjusted goodness of fit measure should be applied for its measurement.

“This adjusted goodness-of-fit measure incorporates the number of observations…as well as the number of independent variables… One can interpret this new measure of fit as penalizing excessive use of independent variables. Instead, one should set up the model as parsimoniously as possible. To take most advantage of the set of possible independent variables, one should consider those that contribute a maximum of explanatory variation to the regression. That is, one has to balance the cost of additional independent variables and reduction in the adjusted R^2." (Fabozzi, et al., 2014)

- Sample size. The number of regressors.

As mentioned by Harrel "in many situations a fitted regression model is likely to be reliable when the number of predictors (or candidate predictors if using variable selection) p is less than m/10 or m/20. ...For "average" subjects, m/10 was adequate for preventing expected errors > 0.1" (Frank E. Harrell, 2015)

Consequently, the number of candidate predictors m/10 (where m – is the sample size) corresponds to 10% of the significance level.

5 Secondary research data collection.

The main source of data for the secondary research for statistical analysis is Yahoo! Finance. It is "a media property that is part of the Yahoo! network that provides financial news, data and commentary including stock quotes, press releases, financial reports, and original content." (Yahoo! Finance, 2022)

The data is collected as cross-sectional data set that is a data set which" consists of a sample of individuals, households, firms, cities, states, countries, or a variety of other units, taken at a given point in time. In a pure cross-sectional analysis, we would ignore any minor timing differences in collecting the data. An important feature of cross-sectional data is that we can often assume that they have been obtained by random sampling from the underlying population." (Wooldridge, 2015)

The particular interest for this study is retrievable financial data and ratios for publicly traded companies for Tesla Inc. and its peers (comparable companies)

(Figure 6). It includes data on market cap, EBITDA, revenue TTM, net income TTM, both revenue and net income growth rates for last year, price to book ratio, forward P/E, EPS current year and EPS next year. The data is relevant for 30.01.2022.

The trailing TTM metrics such as trailing EPS and trailing P/E as well as P/S ratio were intentionally not included in the econometric modelling since they can be presented as linear combinations of the variables already included in the model. For instance, P/S ratio is basically market cap divided by revenue, trailing P/E, on the other hand is market cap divided by net income. The marginal effect of revenue and net income on market cap should be both separate and without multicollinearity associated with those redundant metrics.

n	Symbol	(currency in USD) Company Name	Market Cap (B)	EBITDA (B)	Revenue TTM (B)	Net Income TTM (B)	Revenue growth	Net Income growth	Price/ Book	Forward P/E	EPS current year	EPS next year
1	AAPL	Apple Inc.	2782.134	128.218	378.323	100.555	28.6%	57.3%	38.69	27.47	6.01	5.75
2	MSFT	Microsoft Corporation	2311.046	90.830	184.903	71.185	20.6%	38.7%	14.45	28.81	9.39	9.34
3	TSLA	Tesla, Inc.	849.955	7.166	46.848	3.437	66.3%	518.2%	31.41	69.2	9.99	12.23
4	TSM	Taiwan Semiconductor Manufacturing Company Limited	609.925	38.592	57.112	21.435	18.5%	15.0%	0.28	21.23	3.24	4.34
5	NVDA	NVIDIA Corporation	569.173	9.759	24.274	8.207	64.3%	114.5%	24.01	44.18	7.81	10.52
6	JNJ	Johnson & Johnson	452.254	31.943	93.775	20.878	13.6%	41.9%	6.44	15.59	18.08	21.64
7	UNH	UnitedHealth Group Incorporated	438.959	27.073	285.273	17.285	11.6%	12.2%	5.93	18.86	3.5	39.12
8	PFE	Pfizer Inc.	304.947	27.306	63.600	19.180	30.7%	120.8%	4.03	1	19.48	18.49
9	TM	Toyota Motor Corporation	266.881	41.943	272.490	27.321	20.4%	119.5%	0.01	16.87	15.98	19.55
10	ASML	ASML Holding N.V.	262.164	7.918	21.030	6.423	33.1%	63.1%	25.72	28.35	10.02	13.79
11	ADBE	Adobe Inc.	244.416	6.378	15.785	4.822	22.7%	-8.3%	16.63	31.85	4.2	12.68
12	ABBV	AbbVie Inc.	243.826	28.855	56.197	11.542	22.7%	-74.8%	17.99	9.84	2.69	3.42
13	CSCO	Cisco Systems, Inc.	234.541	15.864	50.789	11.397	5.7%	8.9%	5.49	15.11	6.55	7.25
14	LLY	Eli Lilly and Company	234.461	9.657	28.318	5.582	15.4%	-9.9%	28.66	32.55	15	33.08
15	AVGO	Broadcom Inc.	231.251	14.724	27.450	23.888	14.9%	707.0%	9.27	15.53	21.54	23.69
16	TMO	Thermo Fisher Scientific Inc.	225.407	13.284	39.211	7.725	21.7%	21.2%	5.82	26.12	3.94	4.84
17	ABT	Abbott Laboratories	221.655	12.817	43.075	7.071	24.5%	57.3%	6.44	24.48	2.99	3.13
18	NVO	Novo Nordisk A/S	220.348	7.854	18.304	6.208	10.9%	13.3%	3.35	29.86	1.82	4.68
19	CRM	salesforce.com, inc.	218.798	3.405	24.983	1.739	23.2%	-51.1%	3.83	47.16	9.61	9.41
20	ACN	Accenture plc	216.682	9.030	53.736	6.226	20.1%	18.6%	10.65	32.65	3.54	4.82
21	ORCL	Oracle Corporation	214.918	18.408	41.400	10.263	5.1%	-1.1%	-21.28	15.3	2.83	5.75
22	MRK	Merck & Co., Inc.	204.349	21.057	50.157	47.995	4.5%	579.0%	5.71	11.14	8.62	10.41
23	DHR	Danaher Corporation	200.875	9.633	29.453	7.027	32.2%	92.7%	4.8	24.79	4.86	3.55
24	INTC	Intel Corporation	194.357	33.874	79.024	19.868	1.5%	-4.9%	2.04	12.83	4.33	6.33
25	NVS	Novartis AG	192.188	20.155	52.877	24.021	6.0%	197.6%	3.39	12.98	7.87	10.73
26	QCOM	QUALCOMM Incorporated	187.675	11.371	36.036	9.986	35.0%	48.4%	18.86	14.33	0.54	3.21
27	AZN	AstraZeneca PLC	183.759	6.325	32.816	1.471	27.2%	-41.1%	4.46	15.21	7.36	11.63
28	UPS	United Parcel Service, Inc.	172.307	10.874	97.287	12.890	15.0%	859.8%	14.31	16.38	8.27	9.03
29	TXN	Texas Instruments Incorporated	163.816	9.911	18.344	7.769	26.9%	38.9%	12.29	18.92	9.95	11.54
30	UNP	Union Pacific Corporation	158.103	11.546	21.804	6.523	11.6%	21.9%	11.12	19.2	7.58	11.7
31	INTU	Intuit Inc.	151.443	2.890	10.317	2.092	31.6%	6.4%	15.01	39.12	5.73	8.21
32	CVS	CVS Health Corporation	144.243	17.824	285.061	7.577	7.1%	-4.7%	1.94	13.2	-2.42	7.48
33	BMY	Bristol-Myers Squibb Company	144.077	19.712	45.468	-5.405	15.4%	-12184%	3.87	8.25	5.38	6.99
34	SAP	SAP SE	143.873	6.775	27.842	5.252	1.8%	2.1%	4.09	20.36	7.77	8.04
35	HON	Honeywell International Inc.	139.055	8.730	34.392	5.542	5.4%	16.0%	7.79	22.52	3.47	5.68
36	MDT	Medtronic plc	138.664	9.597	31.797	4.704	14.1%	32.8%	2.67	16.8	2.56	4.87
37	RTX	Raytheon Technologies Corporation	135.204	11.476	64.388	3.864	1.5%	-209.8%	1.84	15.36	6.54	6.03
38	SONY	Sony Group Corporation	132.429	12.824	85.938	75.449	13.9%	704.9%	0.02	15.9	2.73	3.79
39	SNY	Sanofi	132.118	11.272	38.496	37.650	2.2%	235.5%	2.09	12.32	9.7	16.89
40	AMGN	Amgen Inc.	129.067	12.276	25.767	5.609	3.1%	-23.7%	15.76	12.75	3.24	2.64
41	AMD	Advanced Micro Devices, Inc.	127.089	3.423	16.434	3.162	68.3%	27.0%	17.87	31.32	6.35	9.83
42	IBM	International Business Machines Corporation	120.555	11.999	70.787	5.742	-3.9%	2.8%	6.39	12.86	6.4	8.2
43	AMAT	Applied Materials, Inc.	117.615	7.651	23.063	5.888	34.1%	62.7%	9.65	14.87	18.99	22.23
44	DE	Deere & Company	115.201	9.040	43.582	5.964	23.5%	116.7%	6.25	14.85	2.3	3.06

Figure 6. Combined sample of Tesla Inc. and the peer group (30.01.2022) (Yahoo! Finance, 2022).

n	Symbol	(currency in USD) Company Name	Market Cap (B)	EBITDA (B)	Revenue TTM (B)	Net Income TTM (B)	Revenue growth	Net Income growth	Price/ Book	Forward P/E	EPS current year	EPS next year
45	GSK	GlaxoSmithKline plc	113.844	10.230	33.326	4.313	-2.7%	-32.3%	7.35	13.71	-7.15	3.46
46	BA	The Boeing Company	112.230	-0.716	62.286	-4.290	7.1%	-63.9%	-7.48	25.31	1.13	7.36
47	NOW	ServiceNow, Inc.	111.655	0.747	5.895	0.230	30.4%	94.1%	30.22	60.2	27	7.96
48	SHOP	Shopify Inc.	109.757	0.516	4.210	3.410	71.3%	1636.1%	9.63	100.14	9.35	10.38
49	CAT	Caterpillar Inc.	108.816	9.327	50.971	6.489	22.1%	116.4%	6.53	18.4	24.73	28.59
50	ANTM	Anthem, Inc.	107.302	10.046	138.639	6.104	13.8%	33.5%	3.01	13.64	22.76	26.49
51	LMT	Lockheed Martin Corporation	107.065	9.010	67.044	6.315	2.5%	-7.6%	9.72	14.07	-6.16	3.75
52	GE	General Electric Company	101.139	6.677	74.131	-6.519	-6.9%	-214.3%	2.51	16.84	4.66	4.94
53	ISRG	Intuitive Surgical, Inc.	97.093	2.127	5.710	1.705	31.0%	60.7%	8.13	46.46	0.68	0.71
54	INFY	Infosys Limited	94.972	4.084	15.643	2.908	16.3%	16.0%	10.18	27.63	5.21	10.14
55	SYK	Stryker Corporation	93.703	3.677	17.108	1.994	19.2%	24.7%	6.3	22.14	10.12	10.39
56	MMM	3M Company	93.268	9.581	35.355	5.919	9.9%	9.9%	6.17	14.72	4.15	4.67
57	ZTS	Zoetis Inc.	92.402	3.136	7.616	1.980	16.4%	19.1%	19.76	37.56	6.46	8.97
58	MU	Micron Technology, Inc.	88.765	14.989	29.619	7.364	34.2%	145.5%	1.93	6.78	5.84	8.15
59	GILD	Gilead Sciences, Inc.	86.377	14.927	27.305	6.219	10.6%	4956.1%	4.03	9.99	5.44	4.52
60	CNI	Canadian National Railway Company	84.692	8.065	14.477	4.892	4.8%	37.3%	3.69	22.65	6.46	6.82
61	ADP	Automatic Data Proces	83.773	4.094	14.012	2.744	-4.0%	10.5%	16.59	26.32	-2.11	1.62
62	TEAM	Atlassian Corporation Plc	80.671	0.075	2.431	-0.531	34.8%	-55.3%	652.7	152.71	0.71	1.93
63	F	Ford Motor Company	78.087	8.972	134.615	2.867	2.8%	-1859%	2.13	9.58	24.03	20.43
64	CI	Cigna Corporation	77.372	10.155	170.368	8.384	8.5%	58.2%	1.63	10.38	1.68	1.8
65	CSX	CSX Corporation	76.365	6.560	12.522	3.781	18.3%	36.7%	5.69	17.66	5.76	12.62
66	BDX	Becton, Dickinson and Company	76.332	5.267	19.927	1.766	9.4%	10.4%	3.16	19.41	7.47	6.87
67	GM	General Motors Company	72.941	17.820	130.938	11.124	13.1%	228.4%	1.39	7.26	8.6	8.48
68	ITW	Illinois Tool Works Inc.	72.742	3.959	14.455	2.694	15.0%	27.7%	20.83	25.47	3.65	3.17
69	CP	Canadian Pacific Railway Limited	67.609	4.580	7.995	2.852	3.7%	16.7%	5.09	20.89	0.89	1.41
70	ABB	ABB Ltd	66.218	4.365	28.560	1.827	9.8%	-67.1%	4.85	20.58	12.11	13.81
71	NSC	Norfolk Southern Corpo	65.322	5.628	11.142	3.005	13.8%	49.3%	4.79	17.88	18.18	20.68
72	FDX	FedEx Corporation	64.859	11.722	89.552	4.916	19.8%	100.5%	2.6	10.69	5.15	6.55
73	ETN	Eaton Corporation plc	62.457	3.323	19.517	2.068	6.0%	49.1%	3.91	21.12	4.12	4.86
74	WM	Waste Management, Inc.	61.869	4.932	17.931	1.816	17.8%	21.4%	8.63	26.75	4.52	N/A
75	STLA	Stellantis N.V.	60.145	15.856	97.944	0.033	-19.9%	-99.6%	1.3	4.73	43.54	25.07
76	NOC	Northrop Grumman Corporation	59.350	8.734	35.667	7.005	-3.1%	119.7%	4.6	13.72	11.55	12.21
77	GD	General Dynamics Corporation	59.083	5.053	38.469	3.257	1.4%	3%	3.34	15.3	3.82	4.91
78	EMR	Emerson Electric Co.	54.036	3.893	18.548	2.754	10.4%	32.1%	5.47	17.02	12.92	1.98
79	TRI	Thomson Reuters Corporation	51.406	1.564	6.254	6.426	5.1%	241.1%	3.6	42.79	2.27	2.92
80	JCI	Johnson Controls International plc	50.683	4.170	21.660	1.605	-1.9%	73.9%	2.91	20.38	-25.37	-6.65
81	RIVN	Rivian Automotive, Inc.	50.460	-2.028	0.001	-1.233	N/A	249.3%	-1.61	-11.45	4.48	N/A
82	HMC	Honda Motor Co., Ltd.	50.076	18.757	125.137	7.713	10.8%	259.0%	0.01	9.45	-12.57	-1.42
83	LCID	Lucid Group, Inc.	44.699	-1.213	0.000	-1.113	-30.5%	715275%	9.2	-26.11	5.3	4.32
84	RACE	Ferrari N.V.	41.698	1.449	4.710	0.993	25.6%	71.2%	20.33	47.79	-1.02	-0.78
85	NIO	NIO Inc.	33.243	-0.249	5.046	-0.671	153.0%	0.0%	1.36	-90.87	-0.86	-0.91
86	XPEV	XPeng Inc.	27.638	-0.776	2.429	-0.698	317.9%	48.4%	0.62	-41.75	-184.9	N/A
87	TTM	Tata Motors Limited	25.096	3.002	0.032	-0.002	-4.3%	25.2%	0.05	7.06	-0.09	-0.08
88	LI	Li Auto Inc.	24.758	-0.101	3.286	-0.082	286.8%	-17.5%	0.55	-799	-0.08	-0.08

Figure 6. (continuation) (Yahoo! Finance, 2022).

6 Data analysis.

6.1 Econometric modelling.

- Tools for econometric analysis. GRETL software.

As an IT-tool for econometric data analysis, we will use a free open-source software GRETL (Gnu regression, econometrics and time-series library) that is "a cross-platform software package for econometric analysis, written in the C programming language. " (GRETL, 2022)

"Copyright (C) 2007 Free Software Foundation, Inc. <http://fsf.org/>
Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed. " (GRETL, 2022)

- The process of building and modifying econometric models.

We will apply a stepwise exclusion regression method described earlier and, therefore, we start with a specification of the dependent and independent variables to be included in the common model (full model).

According to the main goal of this research we aim to test a market (relative) approach in corporate valuation with the real-world financial valuation data and for that reason, market cap should be selected as an explained (dependant) variable while EBITDA, revenue TTM, net income TTM, revenue growth, net income growth, price/book ratio, forward P/E, EPS current year and EPS next year as explanatory (independent) variables. In this way, we can compute the marginal effect of every independent variable on market cap. Below is the full model in common terms.

[common model]: Mar_Cap = b_0 + b_1* EBITDA + b_2* Rev_TTM + b_3* Net_TTM + b_4*Rev_Gr + b_5* Net_Inc_Gr + b_6* EPS_curr + b_7* EPS_next + b_8* Forward_PE + b_9* Price_Book + ε;

Where "b_0" to "b_9" are regression coefficients and "e" is an error term.

- Setting up significance level.

In implementing the stepwise regression, Fabozzi specifies a "10% significance level for deleting or adding an explanatory variable in the stepwise regression procedure." (Fabozzi, et al., 2014)

In Gretl software ***,** and * indicators mean 1%, 5% and 10% significance level respectively.

Using Gretl we build model 1(below):

Model 1: OLS, using observations 1-87 (n = 82)
Missing or incomplete observations dropped: 5
Dependent variable: Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−44.8734	31.3887	−1.430	0.1572	
EBITDA	21.9651	1.81926	12.07	<0.0001	***
Rev_TTM	−0.738329	0.344140	−2.145	0.0353	**
Net_TTM	0.674923	1.63994	0.4116	0.6819	
Rev_Gr	90.3937	45.2529	1.998	0.0495	**
Net_Inc_Gr	0.0239864	0.0213136	1.125	0.2642	
EPS_curr	1.90146	2.89247	0.6574	0.5130	
EPS_next	−4.04152	2.89762	−1.395	0.1674	
Forward_PE	0.332521	0.228546	1.455	0.1500	
Price_Book	0.164130	0.230212	0.7129	0.4782	

Mean dependent var	207.9843	S.D. dependent var	388.8194
Sum squared resid	1448021	S.E. of regression	141.8147
R-squared	0.881752	Adjusted R-squared	0.866971
F(9, 72)	59.65437	P-value(F)	8.09e-30
Log-likelihood	−517.2915	Akaike criterion	1054.583
Schwarz criterion	1078.650	Hannan-Quinn	1064.246

[Model 1]: ^Mar_Cap = -44.9 + 22.0*EBITDA - 0.738*Rev_TTM + 0.675*Net_TTM + 90.4*Rev_Gr + 0.0240*Net_Inc_Gr + 1.90*EPS_curr - 4.04*EPS_next + 0.333*Forward_PE + 0.164*Price_Book;

As we can see from the p-values of explanatory variables we should modify the model by excluding individually a presumably insignificant variable with the highest p-value, which is Net_TTM (0.6819). In order to reject the null hypothesis of individual significance at 10% significance the corresponding p-value should be higher than 0.10.

Using Gretl we build model 2 (below):

Model 2: OLS, using observations 1-87 (n = 82)
Missing or incomplete observations dropped: 5
Dependent variable: Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−44.5603	31.2004	−1.428	0.1575	
EBITDA	22.4954	1.27682	17.62	<0.0001	***
Rev_TTM	−0.743325	0.341963	−2.174	0.0330	**
Rev_Gr	90.1392	44.9905	2.004	0.0488	**
Net_Inc_Gr	0.0241467	0.0211885	1.140	0.2582	
EPS_curr	2.02270	2.86101	0.7070	0.4818	
EPS_next	−4.14635	2.86994	−1.445	0.1528	
Forward_PE	0.331820	0.227236	1.460	0.1485	
Price_Book	0.163761	0.228897	0.7154	0.4766	

Mean dependent var	207.9843	S.D. dependent var	388.8194
Sum squared resid	1451428	S.E. of regression	141.0056
R-squared	0.881474	Adjusted R-squared	0.868485
F(8, 73)	67.86216	P-value(F)	1.03e-30
Log-likelihood	−517.3879	Akaike criterion	1052.776
Schwarz criterion	1074.436	Hannan-Quinn	1061.472

[Model 2]: ^Mar_Cap = -44.9 + 22.0*EBITDA - 0.738*Rev_TTM + 0.675*Net_TTM + 90.4*Rev_Gr + 0.0240*Net_Inc_Gr + 1.90*EPS_curr - 4.04*EPS_next + 0.333*Forward_PE + 0.164*Price_Book;

Now p-value for EPS_curr is the highest with p-value of 0.4818.

Using Gretl we build model 3 (below):

Model 3: OLS, using observations 1-87 (n = 82)
Missing or incomplete observations dropped: 5
Dependent variable: Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−42.9486	31.0117	−1.385	0.1702	
EBITDA	22.3960	1.26475	17.71	<0.0001	***
Rev_TTM	−0.712580	0.338039	−2.108	0.0384	**
Rev_Gr	91.8502	44.7733	2.051	0.0438	**
Net_Inc_Gr	0.0208196	0.0205893	1.011	0.3152	
EPS_next	−2.78728	2.12379	−1.312	0.1934	
Forward_PE	0.345984	0.225585	1.534	0.1294	
Price_Book	0.148007	0.227039	0.6519	0.5165	

Mean dependent var	207.9843	S.D. dependent var	388.8194
Sum squared resid	1461366	S.E. of regression	140.5282
R-squared	0.880662	Adjusted R-squared	0.869373
F(7, 74)	78.01263	P-value(F)	1.43e-31
Log-likelihood	−517.6676	Akaike criterion	1051.335
Schwarz criterion	1070.589	Hannan-Quinn	1059.065

[Model 3]: ^Mar_Cap = -42.9 + 22.4*EBITDA - 0.713*Rev_TTM + 91.9*Rev_Gr + 0.0208*Net_Inc_Gr - 2.79*EPS_next + 0.346*Forward_PE +
+ 0.148*Price_Book;

This procedure will be repeated step by step for each insignificant variable until we get our final model (Model 8).

Using Gretl we build model 4 (below):

Model 4: OLS, using observations 1-87 (n = 82)
Missing or incomplete observations dropped: 5
Dependent variable: Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−40.7476	30.7090	−1.327	0.1886	
EBITDA	22.3882	1.25984	17.77	<0.0001	***
Rev_TTM	−0.717570	0.336654	−2.131	0.0363	**
Rev_Gr	96.7744	43.9620	2.201	0.0308	**
Net_Inc_Gr	0.0210242	0.0205078	1.025	0.3086	
EPS_next	−2.91024	2.10728	−1.381	0.1714	
Forward_PE	0.384013	0.217074	1.769	0.0810	*

Mean dependent var	207.9843	S.D. dependent var	388.8194
Sum squared resid	1469758	S.E. of regression	139.9885
R-squared	0.879977	Adjusted R-squared	0.870375
F(6, 75)	91.64657	P-value(F)	1.76e-32
Log-likelihood	−517.9024	Akaike criterion	1049.805
Schwarz criterion	1066.652	Hannan-Quinn	1056.569

[Model 4]: ^Mar_Cap = -40.7 + 22.4*EBITDA - 0.718*Rev_TTM + 96.8*Rev_Gr + 0.0210*Net_Inc_Gr - 2.91*EPS_next + 0.384*Forward_PE;

Using Gretl we build model 5 (below):

Model 5: OLS, using observations 1-87 (n = 82)
Missing or incomplete observations dropped: 5
Dependent variable: Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−31.8478	29.4662	−1.081	0.2832	
EBITDA	22.3362	1.25924	17.74	<0.0001	***
Rev_TTM	−0.726738	0.336648	−2.159	0.0340	**
Rev_Gr	86.5535	42.8309	2.021	0.0468	**
EPS_next	−3.25198	2.08144	−1.562	0.1224	
Forward_PE	0.347395	0.214187	1.622	0.1090	

Mean dependent var	207.9843	S.D. dependent var	388.8194
Sum squared resid	1490354	S.E. of regression	140.0354
R-squared	0.878295	Adjusted R-squared	0.870288
F(5, 76)	109.6921	P-value(F)	2.67e-33
Log-likelihood	−518.4730	Akaike criterion	1048.946
Schwarz criterion	1063.386	Hannan-Quinn	1054.744

[Model 5]: ^Mar_Cap = -31.8 + 22.3*EBITDA - 0.73*Rev_TTM + 86.6*Rev_Gr - 3.25*EPS_next + 0.347*Forward_PE;

Using Gretl we build model 6 (below):

Model 6: OLS, using observations 1-87 (n = 82)
Missing or incomplete observations dropped: 5
Dependent variable: Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−59.0210	24.0074	−2.458	0.0162	**
EBITDA	22.4910	1.26703	17.75	<0.0001	***
Rev_TTM	−0.837804	0.332122	−2.523	0.0137	**
Rev_Gr	92.7700	43.0429	2.155	0.0343	**
Forward_PE	0.340621	0.216138	1.576	0.1191	

Mean dependent var	207.9843	S.D. dependent var	388.8194
Sum squared resid	1538222	S.E. of regression	141.3397
R-squared	0.874386	Adjusted R-squared	0.867861
F(4, 77)	133.9972	P-value(F)	7.13e-34
Log-likelihood	−519.7691	Akaike criterion	1049.538
Schwarz criterion	1061.572	Hannan-Quinn	1054.370

[Model 6]: ^Mar_Cap = −59.0 + 22.5*EBITDA - 0.84*Rev_TTM + 92.8*Rev_Gr - 3.25*EPS_next + 0.34*Forward_PE;

Using Gretl we build model 7 (below):

Model 7: OLS, using observations 1-87 (n = 82)
Missing or incomplete observations dropped: 5
Dependent variable: Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−44.4138	22.3556	−1.987	0.0505	*
EBITDA	22.5891	1.27747	17.68	<0.0001	***
Rev_TTM	−0.874303	0.334449	−2.614	0.0107	**
Rev_Gr	50.3395	33.9005	1.485	0.1416	

Mean dependent var	207.9843	S.D. dependent var	388.8194

Sum squared resid	1587837	S.E. of regression	142.6776
R-squared	0.870334	Adjusted R-squared	0.865347
F(3, 78)	174.5157	P-value(F)	1.67e-34
Log-likelihood	−521.0707	Akaike criterion	1050.141
Schwarz criterion	1059.768	Hannan-Quinn	1054.006

[Model 7]: ^Mar_Cap = -44.4 + 22.6*EBITDA - 0.874*Rev_TTM + 50.3*Rev_Gr;

Using Gretl we build model 8 (below):

Model 8: OLS, using observations 1-87 (n = 82)
Missing or incomplete observations dropped: 5
Dependent variable: Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−30.5431	20.4648	−1.492	0.1396	
EBITDA	22.5970	1.28717	17.56	<0.0001	***
Rev_TTM	−0.921121	0.335489	−2.746	0.0075	***

Mean dependent var	207.9843	S.D. dependent var	388.8194
Sum squared resid	1632724	S.E. of regression	143.7616
R-squared	0.866669	Adjusted R-squared	0.863293
F(2, 79)	256.7547	P-value(F)	2.72e-35
Log-likelihood	−522.2136	Akaike criterion	1050.427
Schwarz criterion	1057.647	Hannan-Quinn	1053.326

[Model 8]: ^Mar_Cap = -30.5 + 22.6*EBITDA - 0.921*Rev_TTM;

Finally, we have obtained the model where all the explanatory variables (EBITDA and Rev_TTM) are presumably individually and jointly significant. The last speculation on joint significance is based on p-value for one-sided F-test on joint significance which is around 0. However, in order to make any conclusions about validity of inference we need to check if main assumptions for this regression hold.

- Linearity test.

As mentioned by Montgomery et al., “The assumption of a linear relationship between Y and the regressors is the usual starting point in regression analysis.” (Douglas C. Montgomery, et al., 2012)

Consequently, let’s check the linearity assumption graphically. For that we should draw a graph of residuals against explanatory variables. As can be seen from the graph on Figure 7. there is a systematic pattern in the residuals that depends on EBITDA, other words the scattering is not random which is clear evidence that the relationship between variables is not linear and linearity assumption does not hold, therefore all inference on individual and joint significance is not trustworthy.

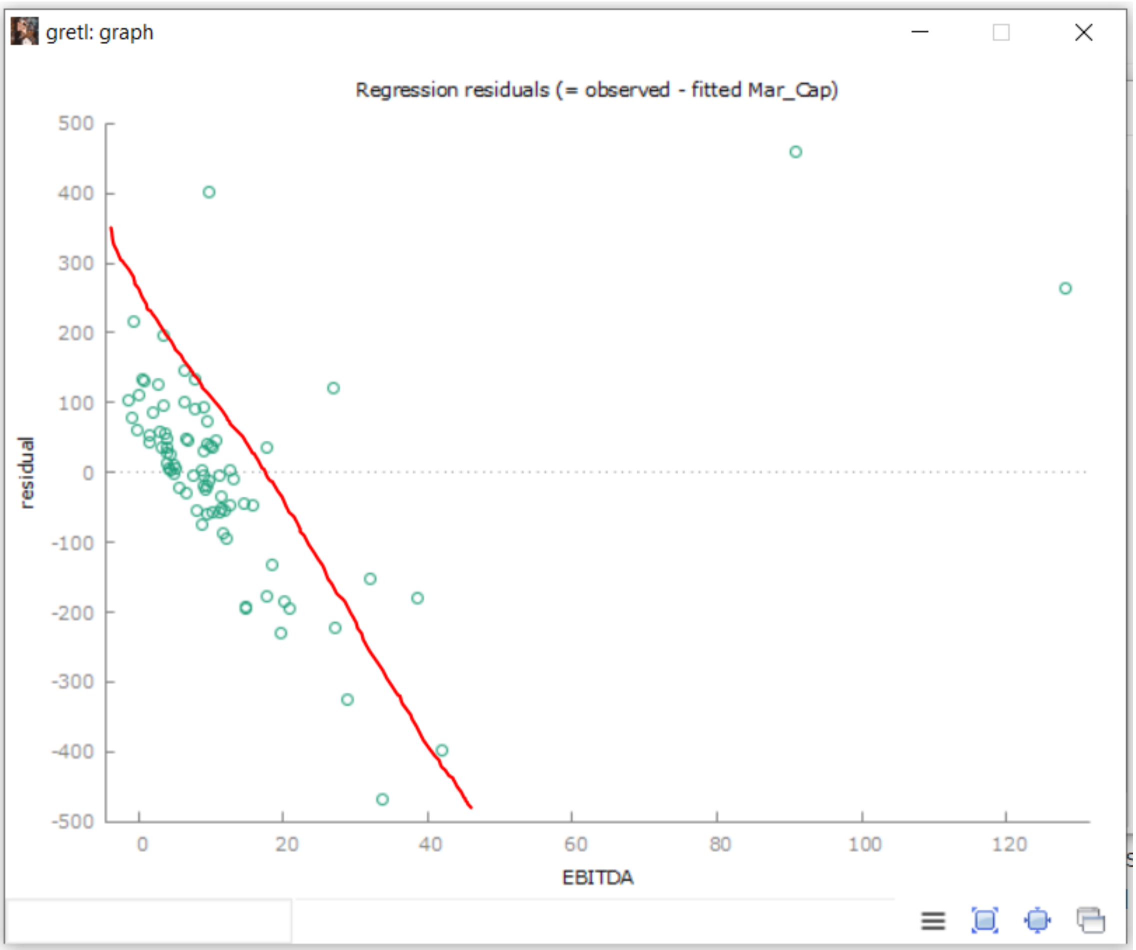

Figure 7. Model 8 regression residuals vs EBITDA (Gretl software).

6.2 Econometric modelling after logarithmic transformation.

"Sometimes prior experience or theoretical considerations may indicate that the relationship between y and the regressors is not linear. In some cases, a nonlinear function can be linearized by using a suitable transformation. Such nonlinear models are called intrinsically or transformable linear." (Douglas C. Montgomery, et al., 2012)

Let's transform our Model 1 (full model) into logarithmic model to make it linear in disturbances. For that we need to put both parts of the equation in natural logarithm form.

[common log-transformed model]: log_Mar_Cap = b_0 + b_1*log_EBITDA +
+ b_2*log_Rev_TTM + b_3*log_Net_TTM + b_4*log_Rev_Gr + b_5*log_Net_Inc_Gr +
+ b_6*log_EPS_curr + b_7*log_EPS_next + b_8*log_Forward_PE +
+ b_9*log_Price_Book + ε;

Where "b_0" to "b_9" are regression coefficients and "ε" is an error term.

- Correlation between independent variables.

Before continuing with modelling let's first analyse the correlation matrix between log_EBITDA, log_Rev_TTM and log_Net_TTM shown on Figure X. The suspicion on correlation between these variables is based upon that all of them can be related due to the nature of how they are calculated. For instance, "EBITDA is essentially net income (or earnings) with interest, taxes, depreciation, and amortization added back" (Investopedia, 2022).

As we can see from the Figure X. there is strong evidence of positive correlation between them. Thus, to separate their marginal effect on log_Mar_Cap we need to breakdown common log-transformed model into next 3 models in every of which is only one of the correlated variables included. Now we can continue modelling.

[common log-model 1]: log_Mar_Cap = b_0 + b_1*log_EBITDA + b_2*log_Rev_Gr + b_3*log_Net_Inc_Gr + b_4*log_EPS_curr + b5*log_EPS_next + + b_6*log_Forward_PE + b_7*log_Price_Book + ε;

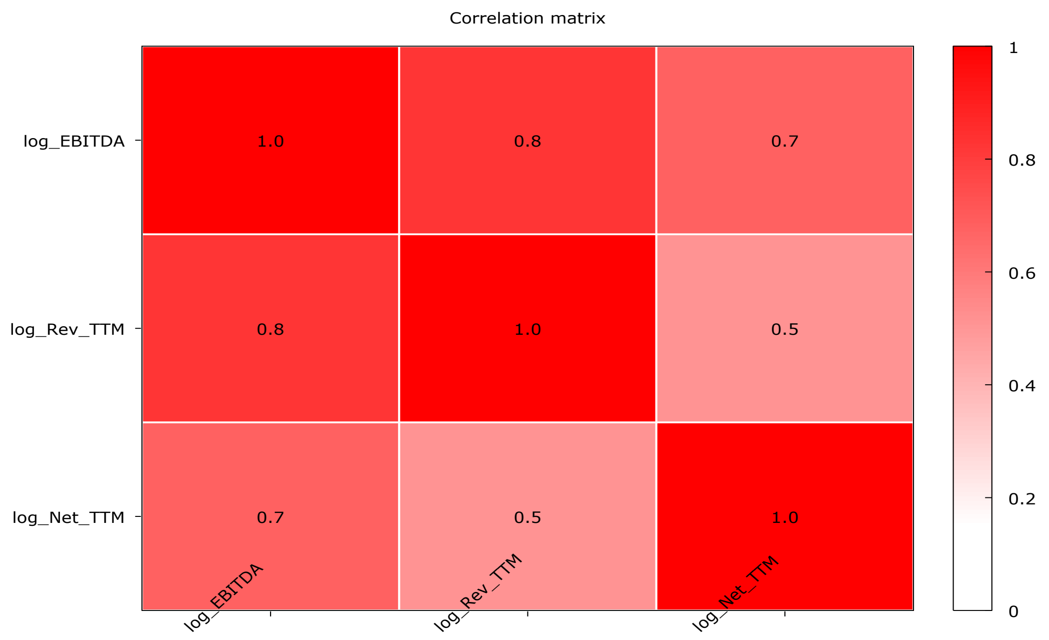

Figure 8. Correlation matrix (Gretl software).

[common log-model 2]: log_Mar_Cap = b_0 + b_1*log_Rev_TTM +
+ b_2*log_Rev_Gr + b_3*log_Net_Inc_Gr + b_4*log_EPS_curr + b_5*log_EPS_next +
+ b_6*log_Forward_PE + b_7*log_Price_Book + ε;

[common log-model 3]: log_Mar_Cap = b_0 + b_1*log_Net_TTM + b_2*log_Rev_Gr
+ b_3*log_Net_Inc_Gr + b_4*log_EPS_curr + b_5*log_EPS_next +
+ b_6*log_Forward_PE + b_7*log_Price_Book + ε;

Using Gretl we build log-models a1, b1 and c1 (below):

Model a1: OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: log_Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−0.647540	0.444771	−1.456	0.1518	
log_EBITDA	0.929648	0.0438469	21.20	<0.0001	***
log_Rev_Gr	0.0264790	0.0511031	0.5181	0.6067	
log_Net_Inc_Gr	0.0263454	0.0257739	1.022	0.3117	
log_EPS_curr	−0.147779	0.0715691	−2.065	0.0443	**
log_EPS_next	0.149193	0.0759924	1.963	0.0553	*
log_Forward_PE	1.17796	0.102851	11.45	<0.0001	***
log_Price_Book	0.0615612	0.0279175	2.205	0.0322	**

Mean dependent var	4.973107	S.D. dependent var	0.821162
Sum squared resid	3.342199	S.E. of regression	0.261167
R-squared	0.911491	Adjusted R-squared	0.898847
F(7, 49)	72.08824	P-value(F)	1.35e-23
Log-likelihood	−0.041459	Akaike criterion	16.08292
Schwarz criterion	32.42733	Hannan-Quinn	22.43491

[model a1]: ^log_Mar_Cap = -0.648 + 0.930*log_EBITDA + 0.0265*log_Rev_Gr + 0.0263*log_Net_Inc_Gr - 0.148*log_EPS_curr + 0.149*log_EPS_next + + 1.18*log_Forward_PE + 0.0616*log_Price_Book;

Model b1: OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: log_Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−0.0715195	0.883937	−0.08091	0.9358	
log_Rev_TTM	0.793215	0.0868768	9.130	<0.0001	***
log_Rev_Gr	0.100441	0.0984436	1.020	0.3126	
log_Net_Inc_Gr	0.0194443	0.0500671	0.3884	0.6994	
log_EPS_curr	−0.307355	0.139523	−2.203	0.0323	**
log_EPS_next	0.0985035	0.148419	0.6637	0.5100	
log_Forward_PE	0.900651	0.194374	4.634	<0.0001	***
log_Price_Book	0.102119	0.0552182	1.849	0.0704	*

Mean dependent var	4.973107	S.D. dependent var	0.821162
Sum squared resid	12.58797	S.E. of regression	0.506850

R-squared	0.666643	Adjusted R-squared	0.619021
F(7, 49)	13.99851	P-value(F)	8.28e-10
Log-likelihood	−37.83567	Akaike criterion	91.67134
Schwarz criterion	108.0158	Hannan-Quinn	98.02334

[model b1]: ^log_Mar_Cap = -0.0715 + 0.793*log_Rev_TTM + 0.100*log_Rev_Gr + 0.0194*log_Net_Inc_Gr - 0.307*log_EPS_curr + 0.0985*log_EPS_next + 0.901*log_Forward_PE + 0.102*log_Price_Book;

Model c1: OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: log_Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	1.53205	0.483963	3.166	0.0027	***
log_Net_TTM	0.774730	0.0459062	16.88	<0.0001	***
log_Rev_Gr	0.227455	0.0613692	3.706	0.0005	***
log_Net_Inc_Gr	−0.169140	0.0322398	−5.246	<0.0001	***
log_EPS_curr	−0.395527	0.0883399	−4.477	<0.0001	***
log_EPS_next	0.406898	0.0927536	4.387	<0.0001	***
log_Forward_PE	0.681476	0.114745	5.939	<0.0001	***
log_Price_Book	0.0937060	0.0344337	2.721	0.0090	***

Mean dependent var	4.973107	S.D. dependent var	0.821162
Sum squared resid	4.991392	S.E. of regression	0.319163
R-squared	0.867817	Adjusted R-squared	0.848934
F(7, 49)	45.95689	P-value(F)	2.22e-19
Log-likelihood	−11.47241	Akaike criterion	38.94482
Schwarz criterion	55.28923	Hannan-Quinn	45.29681

[model c1]: ^log_Mar_Cap = 1.53 + 0.775*log_Net_TTM + 0.227*log_Rev_Gr - 0.169*log_Net_Inc_Gr - 0.396*log_EPS_curr + 0.407*log_EPS_next + + 0.681*log_Forward_PE + 0.0937*log_Price_Book;

The continuation of the stepwise exclusion procedure using Gretl followed earlier gives three end models: model a3, model b4 and model c1.

Model a3: OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: log_Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−0.780908	0.345938	−2.257	0.0283	**
log_EBITDA	0.930731	0.0429835	21.65	<0.0001	***
log_EPS_curr	−0.146417	0.0709166	−2.065	0.0441	**
log_EPS_next	0.159625	0.0736996	2.166	0.0350	**
log_Forward_PE	1.19209	0.0914252	13.04	<0.0001	***
log_Price_Book	0.0597198	0.0274575	2.175	0.0343	**

Mean dependent var	4.973107	S.D. dependent var	0.821162
Sum squared resid	3.472990	S.E. of regression	0.260956
R-squared	0.908028	Adjusted R-squared	0.899011
F(5, 51)	100.7029	P-value(F)	3.39e-25
Log-likelihood	−1.135491	Akaike criterion	14.27098
Schwarz criterion	26.52929	Hannan-Quinn	19.03498

[model a3]: ^log_Mar_Cap = -0.781 + 0.931*log_EBITDA - 0.146*log_EPS_curr + 0.160*log_EPS_next + 1.19*log_Forward_PE + 0.0597*log_Price_Book;

Model b4: OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: log_Mar_Cap

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	−0.431658	0.686683	−0.6286	0.5324	
log_Rev_TTM	0.814972	0.0839296	9.710	<0.0001	***
log_EPS_curr	−0.224485	0.0952897	−2.356	0.0223	**
log_Forward_PE	0.934630	0.168497	5.547	<0.0001	***
log_Price_Book	0.116894	0.0526059	2.222	0.0307	**

Mean dependent var	4.973107	S.D. dependent var	0.821162
Sum squared resid	13.24643	S.E. of regression	0.504717
R-squared	0.649206	Adjusted R-squared	0.622221
F(4, 52)	24.05875	P-value(F)	2.65e-11
Log-likelihood	−39.28878	Akaike criterion	88.57755
Schwarz criterion	98.79281	Hannan-Quinn	92.54755

[model b4]: ^log_Mar_Cap = -0.432 + 0.815*log_Rev_TTM - 0.224*log_EPS_curr + 0.935*log_Forward_PE + 0.117*log_Price_Book;

All the independent variables of these three end models are individually and jointly significant even at the 5% (not to mention 10%) significance level since their correspondent p-values for t-test and F-test are lower than 0.05. However, we cannot yet trust these inferences until all the assumptions for multiple linear regression prove to have been met.

6.3 Assumption checking.

6.3.1 Non-linearity test.

This time we will use formal Ramsey RESET test on misspecification (non-linearity) since, as was mentioned by Prabowo et al., it "is easy to use, it is one of advantages using Ramsey test. While the weakness is that it cannot determine the best alternative model." (Hendri Prabowo, et al., 2020)

Since we are not intending to use this test as a selection tool between models this test should suffice. Therefore, we will use both the squared and cubed terms yhat^2 and yhat^2 consequently.

<u>Ramsey RESET test results for model a3 (Gretl software):</u>

Auxiliary regression for RESET specification test
OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: log_Mar_Cap

	coefficient	std. error	t-ratio	p-value
const	11.4481	10.3673	1.104	0.2749
log_EBITDA	−3.29047	3.65744	−0.8997	0.3727
log_EPS_curr	0.495097	0.574748	0.8614	0.3932
log_EPS_next	−0.547884	0.634944	−0.8629	0.3924
log_Forward_PE	−4.23184	4.70515	−0.8994	0.3728
log_Price_Book	−0.205970	0.227998	−0.9034	0.3707
yhat^2	0.772391	0.695448	1.111	0.2721
yhat^3	−0.0425389	0.0399682	−1.064	0.2924

Test statistic: F = 0.991368,
with p-value = P(F(2,49) > 0.991368) = 0.378

Since P-value is higher than 0.10 we cannot reject null hypothesis at the significance level of 10%. The assumptions that the relationship is linear holds. Thus, the model is correctly specified at least in terms of on linearity.

Ramsey RESET test results for model b4 (Gretl software):

Auxiliary regression for RESET specification test
OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: log_Mar_Cap

	coefficient	std. error	t-ratio	p-value
const	13.7803	16.3974	0.8404	0.4047
log_Rev_TTM	−3.82452	6.03831	−0.6334	0.5294
log_EPS_curr	1.03201	1.65861	0.6222	0.5366
log_Forward_PE	−4.35894	6.91485	−0.6304	0.5313
log_Price_Book	−0.569194	0.864413	−0.6585	0.5133
yhat^2	0.873385	1.36081	0.6418	0.5239
yhat^3	−0.0418144	0.0817073	−0.5118	0.6111

Test statistic: F = 2.378922,

with p-value = P(F(2,50) > 2.37892) = 0.103

For the same reason the assumption on linearity for model b4 also holds.

Ramsey RESET test results for model c1 (Gretl software):

Auxiliary regression for RESET specification test
OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: log_Mar_Cap

	coefficient	std. error	t-ratio	p-value	
const	4.86197	1.33116	3.652	0.0007	***
log_Net_TTM	−5.82440	3.65693	−1.593	0.1179	

log_Rev_Gr	−1.67105	1.06422	−1.570	0.1231
log_Net_Inc_Gr	1.23852	0.787334	1.573	0.1224
log_EPS_curr	3.01599	1.89691	1.590	0.1186
log_EPS_next	−3.10043	1.95388	−1.587	0.1193
log_Forward_PE	−5.14059	3.23121	−1.591	0.1183
log_Price_Book	−0.717191	0.439660	−1.631	0.1095
yhat^2	1.44128	0.868439	1.660	0.1037
yhat^3	−0.0784778	0.0520507	−1.508	0.1383

Test statistic: F = 5.828064,
with p-value = P(F(2,47) > 5.82806) = 0.00548

Since p-value is lower than 0.10 there is enough evidence to reject null hypothesis at the significance level of 10% (and even higher) and conclude that the relationship is non-linear. Since the assumption on linearity does not hold the inference on joint and individual significance as well as prediction power is compromised. Therefore, we will drop this model out of comparison.

6.3.2 Homoscedasticity test.

To check the models for heteroscedasticity we conduct the Breusch-Pagan test.

Breusch Pagan Test results for model a3 (Gretl software):

Breusch-Pagan test for heteroskedasticity
OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: scaled uhat^2

	coefficient	std. error	t-ratio	p-value
const	−0.732448	1.92642	−0.3802	0.7054
log_EBITDA	−0.391426	0.239362	−1.635	0.1081
log_EPS_curr	−0.0256226	0.394913	−0.06488	0.9485
log_EPS_next	0.140878	0.410410	0.3433	0.7328
log_Forward_PE	0.803961	0.509119	1.579	0.1205
log_Price_Book	−0.0635685	0.152902	−0.4157	0.6793

Explained sum of squares = 23.4889

Test statistic: LM = 11.744454,

with p-value = P(Chi-square(5) > 11.744454) = 0.038462

Since p-value is lower than 0.10% there is enough evidence to reject the null hypothesis at the significance level of even 10% and conclude that the homoscedasticity assumption is violated, therefore the inferences can be faulty. Therefore, we drop this model out of the comparison.

Breusch Pagan Test results for model b4 (Gretl software):

Breusch-Pagan test for heteroskedasticity
OLS, using observations 1-87 (n = 57)
Missing or incomplete observations dropped: 30
Dependent variable: scaled uhat^2

	coefficient	std. error	t-ratio	p-value
const	−0 .354977	1.65549	−0.2144	0.8311
log_Rev_TTM	0.333720	0.202341	1.649	0.1051
log_EPS_curr	0 .0885694	0.229729	0.3855	0.7014
log_Forward_PE	−0 .0291753	0.406220	−0.07182	0.9430
log_Price_Book	0.0778433	0.126825	0.6138	0.5420

Explained sum of squares = 7.05823

Test statistic: LM = 3.529117,
with p-value = P(Chi-square(4) > 3.529117) = 0.473465

For this model null hypothesis of homoscedasticity cannot be rejected at 10% significance level due to corresponding p-value higher than 0.10. The assumption holds.

6.3.3 Normality test.

Let's check if the residuals of the model b4 are normally distributed. For that we will use non-normality test provided by Gretl.

Frequency distribution for residual, obs 1-87
number of bins = 7, mean = -9.11552e-016, sd = 0.504717

interval		midpt	frequency	rel.	cum.	
	< -0.76252	-0.93734	3	5.26%	5.26%	*
-0.76252 -	-0.41289	-0.58771	11	19.30%	24.56%	******
-0.41289 -	-0.063255	-0.23807	10	17.54%	42.11%	******
-0.063255-	0.28638	0.11156	17	29.82%	71.93%	**********
0.28638-	0.63601	0.46120	12	21.05%	92.98%	*******
0.63601 -	0.98565	0.81083	3	5.26%	98.25%	*
	>= 0.98565	1.1605	1	1.75%	100.00%	

Missing observations = 30 (34.48%)

Test for null hypothesis of normal distribution:
Chi-square(2) = 0.281 with p-value 0.86895

For this model the null hypothesis that the residuals are normally distributed cannot be rejected at 10% significance level due to corresponding p-value higher than 0.10. Consequently, the assumption about disturbance term holds, therefore inference is valid.

6.3.4 Multicollinearity analysis.

Let's us check if there is a multi-correlation effect within the pool of explanatory variables. For that we will conduct multicollinearity test with Gretl.

Variance Inflation Factors
Minimum possible value = 1.0
Values > 10.0 may indicate a collinearity problem

log_Rev_TTM	1.713
log_EPS_curr	1.041
log_Forward_PE	1.643
log_Price_Book	1.271

VIF(j) = 1/(1 - R(j)^2), where R(j) is the multiple correlation coefficient
between variable j and the other independent variables

Belsley-Kuh-Welsch collinearity diagnostics:

variance proportions

lambda	cond	const	log_Rev_~	log_EPS_~	log_Forw~	log_Pric~
4.417	1.000	0.000	0.002	0.006	0.001	0.012
0.427	3.215	0.000	0.013	0.027	0.000	0.624
0.089	7.064	0.012	0.031	0.922	0.024	0.096
0.061	8.508	0.005	0.452	0.045	0.092	0.265
0.006	27.129	0.982	0.502	0.000	0.883	0.002

lambda = eigenvalues of inverse covariance matrix (smallest is 0.00600168)
cond = condition index
note: variance proportions columns sum to 1.0
According to BKW, cond >= 30 indicates "strong" near linear dependence, and cond between 10 and 30 "moderately strong". Parameter estimates whose variance is mostly associated with problematic cond values may themselves be considered problematic.

Count of condition indices >= 30: 0

Count of condition indices >= 10: 1
Variance proportions >= 0.5 associated with cond >= 10:

const	log_Rev_~	log_Forw~
0.982	0.502	0.883

As we can see from the values for variance inflation factors (VIF) there is no evidence for multicollinearity, therefore the model passed test on multicollinearity.

Since we are working with cross sectional data there is no need to check autocorrelation assumption.

Thus, we have selected the best model of all we had. That is the model b4.

[model b4]: ^log_Mar_Cap = -0.432 + 0.815*log_Rev_TTM - 0.224*log_EPS_curr + 0.935*log_Forward_PE + 0.117*log_Price_Book;

7 Interpretation of results. Tesla Inc. valuation using the model.

According to the number for adjusted R2 this model can explain 62,2% of variations in market capitalization using revenue TTM, EPS current year, forward P/E and price to book ratio as controlled (explanatory) variables. As Moore mentions “if R-squared value 0.5 < r < 0.7 this value is generally considered a moderate effect size”. (Moore, 2013)

Let’s interpret the results and calculate Tesla value derived by the model.

- Ceteris paribus model interpretation.

1% increase in revenue TTM will provoke an 0.82% increase in market cap (keeping all other variables constant).

1% increase in EPS current year will provoke a 0.22% decrease in market cap (keeping all other variables constant).

1% increase in Forward P/E will provoke a 0.94% increase in market cap (keeping all other variables constant).

1% increase in Price/Book will provoke a 0.12% increase in market cap (keeping all other variables constant).

As can be seen in the interpretation above forward P/E closely followed by revenue TTM have the biggest effect on market cap. This can be explained by the fact that investors are looking for solid future earnings but do not forget about past data on revenue. The lowest p-values for these two variables prove this assumption.

- Tesla valuation by the model.

Now we can insert Tesla financial data from Figure X. into the equation.

Ln(Mar_Cap) = -0.432 + 0.815*ln(46.848) - 0.224*ln(9.9) + 0.935*ln(69.2) + 0.117*ln(31.41);

Mar_Cap = EXP(-0.432 + 0.815*ln(46.848) - 0.224*ln(9.9) + 0.935*ln(69.2) + 0.117*ln(31.41));

Mar_Cap = 701,054B ($)

8 Conclusions.

The model shows that Tesla Inc. is overvalued by about USD 150 billion which along makes up the market cap of Ford Motor and General Motors companies combined. Although, this practical application of the econometric modelling to market (relative) corporate valuation shows a rather surprising result it does not necessarily mean that all the companies in the peer group are valued correctly.

Hooke seems to agree with that statement: “The main problem with the comparable company approach is that it doesn’t tell you whether the industry as a whole is cheap or expensive at any specific time. Some practitioners look back to historical norms to identify clear aberrations, but staying with this idea requires a contrarian view that endangers one’s career prospects.” (Jeffrey C. Hooke, 2010)

In addition, it is worth mentioning that during the log model transformation 30 observations which had a negative value were lost which most likely influenced the prediction power of the model by reducing the sample. A rather limited sample size associated with this approach present another problem. “A second problem with relative analysis is the lack of true comparables. Even within the same industry, firms have different characteristics that limit the relevance of such studies.” (Jeffrey C. Hooke, 2010)

In general, this method can be successfully used in conjunction with other methods of corporate valuation giving the valuer/analyst another perspective on the matter.

As a conclusion, Hooke recommends “applying the three alternative approaches as a reality check for every relative valuation. If the calculations are significantly different, the analyst should refrain from making a recommendation until the matter is resolved.” (Jeffrey C. Hooke, 2010)

References

Authers, J., 2021. *Bloomberg Opinion.* [Online]
Available at: https://www.bloomberg.com/opinion/articles/2021-10-27
[Accessed 25 November 2021].

Cassidy, J., 2021. *www.newyorker.com.* [Online]
Available at: https://www.newyorker.com/news/our-columnists/elon-musk-and-the-dangers-of-another-stock-bubble
[Accessed 25 November 2021].

Damodaran, A., 2014. *Applied corporate finance.* 4th edition ed. NJ: Wiley.

Dimitrios Asteriou & Stephen G. Hall, 2007. *Applied econometrics. A modern approach..* Revised Edition ed. New York: PALGRAVE MACMILLAN.

Douglas C. Montgomery, Elizabeth A. Peck & G. Geoffrey Vining, 2012. *Introduction to linear regression analysis.* 5th ed. ed. New Jersey: John Wiley & Sons, Inc..

Dunn, J., 2016. *Insider.* [Online]
Available at: https://www.businessinsider.com/
[Accessed 3 April 2022].

Fabozzi, F. J., Focardi, S. M., Rachev, S. T. & Arshanapalli, B. G., 2014. *The Basics of Financial Econometrics. Tools, Concepts, and Asset Management Applications..* New Jersey: John Wiley & Sons, Inc..

Fazzini, M., 2018. *Business Valuation: Theory and Practice.* 1st ed ed. s.l.:Palgrave Macmillan.

Frank E. Harrell, J., 2015. *Regression Modeling with Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis..* 2nd ed. New York: Springer.

GRETL, 2022. *GRETL.* [Online]
Available at: http://gretl.sourceforge.net/
[Accessed 10 February 2022].

Gujarati, D. N., 2004. *Basic Ecinometrics.* 4th ed. New York: The McGraw−Hill .

Hendri Prabowo, Suhartono Suhartono & Dedy Prastyo, 2020. *The Performance of Ramsey Test, White Test and Terasvirta Test in Detecting Nonlinearity..* [Online]
Available at: https://www.researchgate.net/
[Accessed 7 April 2022].

Herath, S., 2021. *Analytics Vidhya.* [Online]
Available at: https://www.analyticsvidhya.com/
[Accessed 15 April 2022].

IANS, 2021. *Business Standard.* [Online]
Available at: https://www.business-standard.com
[Accessed 3 April 2022].

Investopedia, 2022. *Investopedia.* [Online]
Available at: https://www.investopedia.com/
[Accessed 4 April 2022].

IVSC, 2022. *https://www.rics.org.* [Online]
Available at:
https://www.rics.org/contentassets/542170a3807548a28aebb053152f1c24/ivsc-

effective-31-jan-2022.pdf
[Accessed 3 April 2022].

IVSC, 2022. *IVSC.* [Online]
Available at: https://www.ivsc.org/about/
[Accessed 03 April 2022].

Jeffrey C. Hooke, 2010. *Security analysis and business valuation on Wall Street : a comprehensive guide to today's valuation methods..* 2nd ed. New Jersey: Wiley & Sons.

John Black, Nigar Hashimzade & Gareth Myles, 2009. *A Dictionary of Economics.* 3 ed. ed. s.l.:Oxford University Press.

Lee C. Adkins, 2011. *Using gretl for Principles of Econometrics..* 4th ed. s.l.:Lee C. Adkins..

Lee, K., 2020. *American Century Investments.* [Online]
Available at: https://www.americancentury.com/insights/the-ongoing-rise-of-tesla/
[Accessed 25 April 2022].

Liberto, D., 2021. *Investopedia.* [Online]
Available at: https://www.investopedia.com/
[Accessed 17 April 2022].

Broekema Marc J. R., Niek Strohmaier, Jan A. A. Adriaanse & Jean-Pierre I., 2020. Are Business Valuators Biased? A Psychological Perspective on the Causes of Valuation Disputes.. *Journal of Behavioral Finance,* p. 20.

Moore, D. S. N. W. I. &. F. M. A., 2013. *The basic practice of statistics.* 6th ed. ed. New York: Freeman and Company.

Seth Bernstrom, 2014. *Valuation: The Market Approach.* 1st edition ed. s.l.:Wiley.

Singhi, R., 2021. *Yahoo!Finance.* [Online]
Available at: https://finance.yahoo.com
[Accessed 25 November 2021].

Subramanian, P., 2021. *yahoo!finance.* [Online]
Available at: https://finance.yahoo.com
[Accessed 3 April 2022].

Ullah, M. I., 2020. *Basic Statistics and Data Analysis..* [Online]
Available at: https://itfeature.com/
[Accessed 23 April 2022].

Wigglesworth, R., 2021. *Financial Times.* [Online]
Available at: https://www.ft.com
[Accessed 15 April 2022].

Wooldridge, J. M., 2015. *Introductory Econometrics: A Modern Approach.* 6th edition ed. s.l.:Cengage Learning.

Yahoo! Finance, 2022. *Yahoo! Finance.* [Online]
Available at: https://finance.yahoo.com/
[Accessed 10 February 2022].