UNIVERSITY OF TAMPERE
School of Management

Economic forecasting in a business environment: an OLS estimator application

Case Kalmar

Economics
Master's Thesis
May 2018
Instructor: Prof. Hannu Laurila
Perttu Pärssinen

Abstract

University of Tampere

Master's Programme in Business Studies

PÄRSSINEN, PERTTU: Economic forecasting in a business environment: an OLS estimator application

Master’s Thesis: 64 pages, 6 appendix pages

May 2018

Keywords: Financial forecasting, budgeting, forecast accuracy, OLS estimation,

The practice of financial forecasting has been in the interest of researchers since the late 1970s. Despite highly sophisticated models and increasing competence in econometrics and economics studies, actual business environment has overlooked statistical methods in forecasting. This thesis seeks to bring the usefulness of econometrical studies to business environment and for finance organizations’ budgeting processes. The thesis starts with introducing the complexity of forecasting practice in business organizations and the contradicting desires and incentives of different stakeholders. Goal for the empirical part of thesis is to create an econometric model by utilizing OLS estimator for Kalmar forklift trucks sold in geographical area consisting Europe, Middle East and Africa. In the later part of thesis, this model is extended to a forecasting model and the performance of it is evaluated against other forecasts by operations. At the end, the caveat of cyclic sales is analyzed using dummy variables and remarks for the future are denoted.

Our key finding is that by using external lagged variables one can create a fundamental fact based model, which can be used as a highly accurate forecasting model. Using simple OLS regression and common-sense variable, the forecast model can track the actual sales development over the time from year to another. Forecast model has caveat what comes to a human factor. The quarter-oriented economy will influence the revenue recognition process and will make the sales to deviate from its fundamental value.

The forecast model do perform as the literature implies, a simple forecast model can predict sales accurately and most importantly, fact based. The forecast model would bring value to the complexity of budgeting and rolling forecasting, since it would bring non-biased forecast and on top of which one can build a complete financial plan. When using the forecast model management would be able to take calculated risks based on facts and selected risk level. The idea and concept, which was proven in thesis could and should be extended to cover entire Kalmar mobile equipment division.

Tiivistelmä

Tampereen yliopisto

Kauppatieteiden maisteriopinnot

PÄRSSINEN, PERTTU: Economic forecasting in a business environment: an OLS estimator application

Pro Gradu -tutkielma: 64 sivua, 6 liitesivua

Toukokuu 2018

Avainsanat: Ennustaminen, budjetointi, ennustetarkkuus, OLS-regressio,

Liikeyritysten taloudellinen ennustaminen on ollut tutkijoiden mielenkiinnon kohteena 1970-luvun loppupuolelta lähtien. Tutkijat ovat etsineet parhaita käytäntöjä, joita yritykset käyttävät taloudelliseen ennustamiseen. Ekonometrinen tai tilastotieteellinen ennustaminen ei ole yritysten suosima ennustetapa, vaikkakin mallit ovat kehittyneet sekä niiden käyttöönotto on helpottunut. Matemaattisia ennustemalleja on ylenkatsottu ja katsotaan edelleen, vaikka niiden puolueettomuus sekä ennustetarkkuus ovat parempia kuin ihmisten intuition ja kokemukseen perustuvien ennusteiden on.

Tämän tutkimuksen tarkoituksena on tuoda esiin ekonometristen mallien hyödyllisyys liike-elämään sekä taloushallinnon budjetointiprosessiin. Työ alkaa ennustekäytäntöjen tarkasteluilla jo tehtyjen tutkimusten perusteella sekä avaa ristiriistaisten insentiivien vaikutussuhteita ennusteprosessiin sekä budjetointiin. Työn empiirisessä osassa mallinnetaan OLS-regressiolla Kalmarin haarukkatrukkien myyntiä maantieteellisellä alueella, joka koostuu Euroopasta, Lähi-idästä sekä Afrikasta. Estimoitu malli laajennetaan ennustemalliksi, jonka ennustetarkkuutta verrataan ja arvioidaan muihin taloushallinnon tekemiin ennusteisiin. Lopuksi käsitellään mallin heikkouksia sekä mahdollisia tulevaisuuden mallinnustapoja, joilla myynnin syklisyyttä voitaisiin paremmin mallintaa.

Työn tärkein havainto on mahdollisuus luoda täysin ulkoisiin fundamentaalisiin faktoihin perustuva ennustemalli. Yksinkertainen OLS-regressio yhdistettynä fundamenttimuuttujiin mahdollistaa suuren ennustetarkkuuden tilikaudesta toiseen. Ennustemalli toimii kuten aiempi tutkimus sekä kirjallisuus osoittaa. Ennustemalli tuottaa objektiivisen ennusteen myynnin kehittymisestä, kun myyntiin ei kohdistu ei-fundamentaalisia vaikutuksia kuten sisäisiä insentiivejä myynnintulouttamisen suhteen. Ennustemalli tuo selkeyttä budjetointiin ja antaa selkeän suunnan myynnin kehittymiselle. Ennustemallin hyvänä puolena mainittakoon siitä saatava ymmärrys myynnin todennäköisyyksille, joka mahdollistaa johdon harkitun riskinoton. Mallin idea ja konsepti on todistettu ja seuraava vaihe on laajentaa se kattamaan koko Kalmarin mobiilikonedivisioona.

Contents

List of figures

List of tables

List of appendices

1 Introduction

Current financial life and its many factors make these days the best to experience. Currently interest rates are at an all-time low while stock valuations are high. Stock indices have achieved all-time highs repeatedly; investing has become trendy and young people have found their way to stock market. Increasing demand from public and exceptional financial times could and should mean that the companies listed in the stock market need to be perform better every day. There are new financial analysis providers that operate in a fast pace and provide information for the PlayStation-genre, e.g., Inderes Ltd. in Finland. These services use social media efficiently, fast and most importantly, they reach an entirely new audience with their disruptive content.

Increasing publicity and demand for accurate information should force companies to improve their budgeting and forecasting processes. When market and the audience demand better and better results in every interim report, there is not much room for negative profit warnings or disappointments. When valuations are high, drops can be large and one can see this easily in stock exchanges at the end of each quarter. Market will punish with a fast and clear signal when the performance is not adequate. For example Nokia Corp. 26.10.2017 in Nasdaq OMX Helsinki, drop in a single business day was 18%. Most likely, it was not a great day to be the Nokia's Chairman of the board when the market wiped off almost one fifth of the market value of the company.

There are few question for the research to which efforts to find plausible answers and investigations are focused on. The purpose of the thesis and research is to look for external factors that drive the sales and order intake of Kalmar mobile equipment. Macroeconomy and global trade will have an impact to sales and order intake but whether it is possible to identify those factors exactly and statistically significantly is the intriguing question. The search for explaining factors can be divided to different categories: financial factors, industry specific factors and global factors.

With industry specific factors is meant for example steel and concrete industries that should have high correlation with forklift truck sales since forklift trucks are heavily used in those industrial fields. The key question is whether one can derive an external facts based econometric model to explain past history of external sales and if such is found, can one forecast by using it. Such model should be possible to find but not with ease. Ideally, the external fact based model would use macroeconomic variables and financial markets to explain the past. Based on conversations, forecast model that would show the direction of sales with high enough confidence would be highly useful for budgeting.

Macroeconomic variables will be business and product line related. Financial market related variables could be common between different product lines at least to some extent.

Highly interesting question is the effect of financial markets to sales whether one can find relationship between financial markets and the sales. Financial factors include all financial market factors and financial policies made by central banks. Financial markets have been highly volatile and irruptive past years; one will remember the financial crisis from 2009 onward and the Greek government debt crisis. Since those days, operating conditions have become much more favorable and currently cost of debt for good investment grade companies is virtually free.

Investments using leverage are now highly compelling, almost every possibility to invest that has higher cash flow than the initial investment, will have a positive net present value and will add value to company and to its shareholders. If one cannot find a link between sales and financial markets, one could deduce that investment decisions are not made with real cash flow but instead with nominal rates. The question then is, whether the machine is purchased at a specific moment, only because it can be fitted to budget or because it is more profitable now. This link can be quite different between customers. Large customers are more likely to use investment calculations for investment decision making. Small customers on the other hand might not have the chance to wait for the best time to investment and they will purchase the machine when needed.

If one can find a correlation, or more preferably a causal relationship, between macroeconomic variables and sales, how much further can be forecasted? One key question then is what the accuracy for the forecast would be. Ideally, the model derived should be transformed to a medium-term forecast model with approximately one-year forecasting capability. If the forecast horizon is short, the usage of the forecast is not meaningful. For static budgeting process, the forecast should be a bit longer, from 15 to 18 months but for rolling forecasting shorter one-year period is sufficient.

Goal for the research is to create a model that can forecast nearly one year ahead. Naturally, when one has estimated a model that forecasts for example one year ahead, the question arises: “We would prefer forecast for a year and a half forward”. Naturally, the one-year forecast model will not do that and then one must take an opinion about the development of the fundamental variables to have enough data points to perform the forecast. In such case, where the underlying facts are forecasts, will not be the best possible solution. One can perform such act, but to the results, one should interpret with some suspicious.

2 Business and research environment

2.1 Budgeting and forecasting in business context

The earlier example about Nokia Corporate's share price drop leads to another important topic: how to budget, forecast and communicate those to market and to public. In order to have something to communicate, one must have first an opinion how the market and business is developing.

By definition a budget is a numerical plan for a company to plan its actions for the future in controllably way. A budget is one of cornerstones for a company and creating an accurate one can be challenging, especially if the budget is created without a forecast model. One can put numbers into spreadsheet, but the accuracy of the budget will be more or less vague. There are many ways to derive budgets but there are some more sophisticated models for budgeting and forecasting. Intuition and experience can be one tool for budgeting and forecasting but how much one should count and trust in numbers derived in this way. Author at least would not too much.

Budgeting should start with defining long-term goals and strategy for a company based on its vision. After a company has defined its long-term position, it is possible to derive budgets for shorter periods and to plan business actions. (Shim, Siegel & Shim 2012, 1-2.) At least there could be three types of budgets, target state for the company: how much we would like to sell in the market. After the target-state has been determined, then one could progress towards costs to sell that amount. The sales target could be for example, increase sales by 15 % and adjust resources accordingly. Second type of budget could be to calculate and determinate all costs associated with a business plan created for the next accounting period. Then calculate sales based on the cost base determined together with a required profit for capital. The required profit can be derived from the required rate of return for capital. The second way to form a budget is called bottom-up method. Third way could be quite similar to the first but only with organic growth, for example with 5 % sales increase and then one will calculate the costs, which are realized to achieve this sales target. (Shim et al., 2012, 1-14.) The importance of the forecasting rises if a company is an industrial company, since those will be always more affected by recessions than consumer product companies will (Makridakis, Wheelwright & Hyndman 1998, 556).

Ideally when performing budgeting, there would be a forecast model to point out the direction of the market and the market share for the company (Makridakis et. al 1998, 505). After the market development has been estimated, base sales forecast could be created. This base forecast would imply a sales level assuming that there will not be any internal or external issues, for example issues to deliver equipment to customers. If the actual sales would differ from the result of the model, then it should be analyzed was it caused by an internal factor or by an external one. An internal factor would be for example incapability to deliver goods to customers. An external one could be high demand boom due to one-time tax policy in a particular country or area. (Shim et al., 2012, 31.) Another important factor on the behalf of statistical model is its objectivity. Operations and management will have conflicting interests and there are political issues with-in a company, which will highly decrease the forecast accuracy. When the human influence is eliminated, an expert's forecast can actually outperform a simple forecast model but only if the humanly bias is eliminated. (Walker & McClelland 1991, 379-381.)

Different stakeholders in a company will opt for different budget figures, marketing would prefer higher whereas executives of production lower (Makridakis et al., 1998, 505). A forecast model created from external factors would give a baseline with a certain probability for the budget. This would be a better way to create a budget and would help the company to resource and focus its actions. Another positive aspect on the behalf of forecast model is the elimination of anchoring (Makridakis et al., 1998, 505). When there is a judgmental bias to make the forecast or the budget to deviate from the value, which is realistic to achieve, is that phenomena called anchoring. Anchoring is more drastic and possibly dangerous and costly if anchoring level comes from the top of the company, from the CEO or from the Chairman of the board. In such cases, the organization will not be willing to deviate from the desires of the highest-ranking person and the forecast or the budget will deviate from the fundamental value drastically. (Makridakis et al., 1998, 505.)

Makridakis et al., advice already in (1998) that organization that are not using statistical methods to forecast should start it as soon as possible. That was roughly twenty years ago and there are still companies that forecast using judgmental decisions and intuition. Makridakis et al., (1998) even noted that the usage of statistical forecasting tools has become a competitive requirement, i.e., not a way to gain competitive advantage over others but a must-have function of a competitive company. Similarly, they noted about extrapolative methods that those would not bring strategic advantage to a company, even though the forecast would be highly accurate, since everyone else can employ the same tools (Makridakis et al., 1998, 567).

For example Shim et al., (2012) advice regression analysis for sales forecasting, which could be highly useful for budgeting. In regression analysis one searches statistically significant dependence relationship between two or more variables. One could formulate a linear or non-linear relationship between the variables. When one has formulated a model that is statistically significant, then a forecast model is possible to create. With a forecast model, it would be possible to create a baseline sales figure for the target period with a distribution for the forecast error. Idea to use forecasting model is to have transparent, objective, solid and a systematic way to forecast. (Makridakis et. al 1998, 505.) However, the forecast model do not require being highly complex, since there is no empirical evidence to support the assumption that a complex and highly mathematical model would consistently outperform a more simple model (Makridakis et. al 1998, 562).

Since forecast is only a forecast, there is uncertainty related to the forecasted value, i.e., a possibility that the sales will deviate from the forecast. Ideally the probability distribution of the forecast error is small and known, in order to motivate the management to approve the usage of statistic forecast models as a base for the budgeting process. Many have argument on the behalf of demand and economy driven models, e.g., (Chase, 2013; Gillilan, Sglavo & Tashman. 2015). On the behalf of statistical forecast model created from economic factors, speaks the common practice to publish financial guidance for market. For example, Cargotec Plc. expected sales for 2016 were announced to be EUR 3,729 million (Cargotec 2016). Competitor of Cargotec Plc., Konecranes Plc. announced its sales to be close to 2016 level in 2017, i.e., 3,278 million EUR (Konecranes 2017). From other field of mechanical engineering Valmet Plc. revised its guidance for 2017 sales in 12.4.2017; their sales will be higher than in 2016 (Valmet 2017). From these publicly stated targets and guidance for sales, one should have a thought beforehand, what exactly to say to public.

First, a company must have a clear and probable sales volume figure to give and it would be highly beneficial if that figure will be achieved. Fascinating question is not only, how the financial guidance is derived, but also, whether it is the same, which the company has budgeted for themselves. Does the management believe in their budget figures enough to present those to the market or is the budget figures larger and then the management makes their own adjustment before communicating to the market? If the budget and publicly communicated values are the same, then it will be highly relevant that the budget figure will be derived from the real economy and not based on intuition or experience (Makridakis et al., 1998, 491). If the budget figure is the same as is the financial guidance, is the financial figure overly boosted to create new appetite for investor to invest to the company (Makridakis et al., 1998)?

If a company communicates large increase in (profitable) sales for the next year, then the normal reaction should be that there are even more investors willing to invest in that particular company and the stock price should increase. If the financial guidance is not the same as the budget is but is lower, then there is the question to ask, did the management play safe and publish lower guidance than they are expecting and targeting? Playing safe should not be attractive since there are always competitors in the market and investors will benchmark similar companies against each other. One's low guidance might make another company more attractive to invest. In addition, what if the company suddenly stops publishing top line guidance.

Another important issue with forecasting is the insignificance of the practice. Forecasting is seen as an addition to other tasks for a group controller or for a market planner on one's way to more important positions with-in a company and on one's career. (Makridakis et al., 1998, 562.) In addition, to highlight the issue of judgmental decisions, different persons will judge same change differently. Makridakis et al., (1998) found that the same factor would increase sales by some participants while others will interpret the factor to cause decrease in sales. Walker and McClelland (1991) found in their study that sales organization has the highest forecast error, followed by finance but surprisingly production had the best accuracy. From this point of view, it is important to forecast but also who is making the forecast, since finance has the most knowledge of the financial performance of the company, i.e., sales, production, costs and so on, but do not have the best forecasting accuracy (Walker & McClelland 1991, 377).

One could implement a procedure for an organization, where a forecast model would create a baseline forecast, and if one wants to deviate from it; all modifications should be presented with arguments and facts (Makridakis et al., 1998, 506). Things might have changed but Meehl (1954) studied decisions makers' behavior and found out that people are inconsistent with their choices made even though the possibilities to select are the same. Similar study has been done on the human behavior, which would imply the validity of Meehl's point. Thaler and Johnson (1990) performed a remarkable study about decision making under uncertainty when there has been a prior event, i.e., either a gain or a loss. The study by Thaler and Johnson (1990) was about the impact of the prior event to the decision to be made. Meehl (1954) suggested to use decisions making rules as the baseline because those would be consistent, the opposite of human behavior and based on Thaler and Johnson (1990), author would recommend this as well.

There should be a forecast model in every firm, since the poor performance of the previously mentioned sales organization was because of salespersons' way to forecast using their intuition and "best guess". In addition to the initial forecast, the salespersons did not have true accountability for their forecasts since when those were reviewed; adjustments were made and agreed by "gut feelings". The top performing organization of that firm, the production, performed forecast on weekly trend levels together with seasonality factors. From these they derived a mathematical model for weekly production volume. The finance organization, with accuracy between production and sales, forecasted using the values derived by production and sales personnel. It is fascinating that the finance did not have enough confidence to either of two other organizations but decided to combine both forecast and create their own. Naturally when mixing great with bad, the result will not be the best model of them all, but something between those or possibly something completely random. (Walker & McClelland 1991, 378.)

Key point from this debate is that, could the statistical forecast model provide more objective and truthful sales figure to use in budgeting and correspondingly to communicate financial guidance to market. At least research implies so (Makridakis et al., 1998; Walker & McClelland 1991). The top management might still lower the communicated value for their own risk buffer but at least the base figure would be fundamentally durable. Questions about how financial guidance is derived will most likely be unanswered to some extent since companies rarely communicate these to public precisely.

2.2 Relationship between macroeconomy and equipment sales

One should have noticed the highly unusual financial and macroeconomic environment since the 2008 financial crisis. Monetary policy by European Central Bank and Federal Reserve has had an effect to interest rates of government bonds and during the latter part of quantitative easing also to corporate bonds. Idea of the quantitative easing is and has been to boost the economy and start the growth process again. By purchasing government debt, central bank is increasing prices of the bonds and correspondingly lowering yields of bonds. Lower yield implies lower costs of debt. When the secure options have low rates of returns, i.e., when the government bonds are no longer a good investment, investors will look for another options. Increasing demand for stocks and corporate bonds increases the prices of those assets and in the case of corporate bond lowers the yield of bonds issued.

New technologies often improve efficiency of new equipment, which will have lower cost per move or per other measured unit. There are two elements in each investment that a firm makes. Fixed cost is the sunk cost that a firm can no longer affect after the investment good has been acquired. Variable cost is a cost element that depends on the usage of the initial investment. In Kalmar’s context, a customer firm’s variable cost is to operate the machine purchased and the variable cost can be counted on, e.g., per move basis or by lifted tons.

New machine can be a good investment in many ways: it can be for example more economical, quieter and less polluting. With these properties, the customer could operate in new areas and for longer times, extend the usage of the same personnel, and the fleet in the same destination. A new machine could also be faster and safer; more work can be done for the same time compared to previous. All of these combined are beneficial for an operative firm and can be calculated. When firm calculates the benefits and costs of the equipment, they will derive net profit for the investment. The net profit can be due to higher revenue with same variable cost as before, same revenue as before but with lower variable cost or a combination of those. Company can then make calculations for the investment profitability using payback period or Net Present Value (equation 1).

Lower yield creates new possibilities for the company to perform investments that have not been profitable in the past. Lower yield of the bonds already issued do not affect the issuer, i.e., the company, but will be beneficial if the company issues new bonds. For example, a 100 000 € initial investment for ten years with a yearly net revenue of 15 000 € and zero residual value. Payback period is 6 years and 8 months. Investment is barely profitable to perform with an 8% rate of return for capital when calculated using the Net Present Value (NPV) (equation 1). The tax benefit of interest rate deductibility is excluded from the analysis. Profitability results for the investment is visualized in the figure 1 with scenarios for different discount rates.

NPV can be calculated in the following way:

$$NPV_t = -CF_0 + \sum_{1}^{s} \frac{CF_{t+s}}{(1+r)^s}, \tag{1}$$

where CF is net cash flow for the period, r is the rate of return required for capital and t stands for current period and the s future periods. In (equation 1) is the presented normal NPV formula, e.g., Brealey (2014) with modifications introduced and argued by Samuelson (1973).

Current yield for an investment grade bond can be as low as one to two percent. Large publicly listed company Metso Corp. issued a 300 million euro bond with 1.125% yield (Metso 2017). When comparing this to figure 1 below, one could see the upper mentioned investment to be highly profitable when discount factor is equal to the interest rate for debt. Low interest rate is one of key inspirations for the thesis and for the econometrical business analysis to be performed. On top of this cost of debt should be added a required profit for equity to the shareholders.

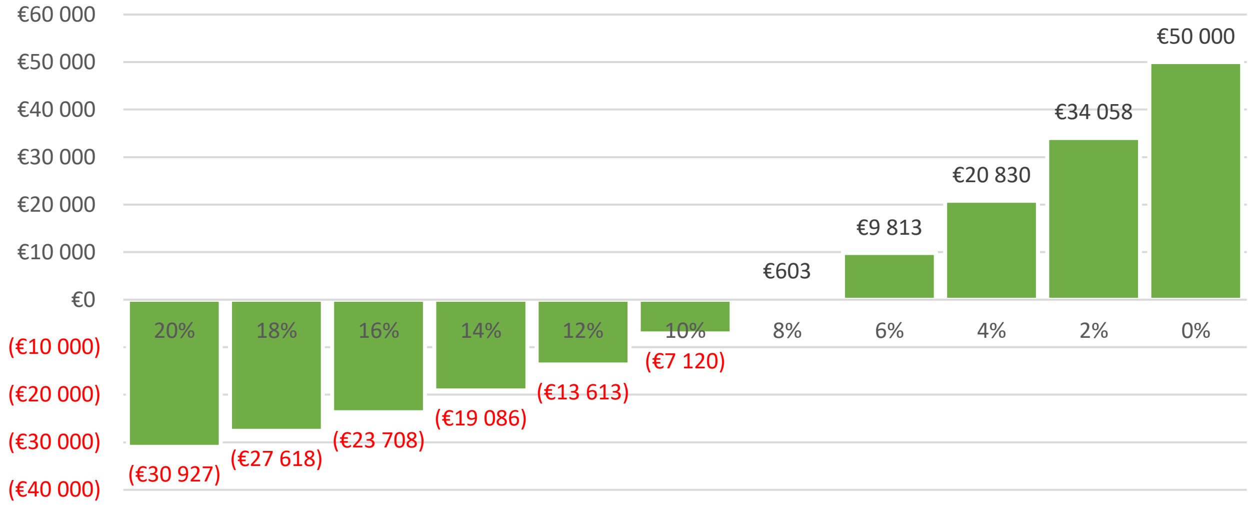


Figure 1. Net Present Value calculation using different discount rates (estimated by the author).

Based on the NPV analysis above in the example and in figure 1, one might conclude that interest rates should be considered in the econometrical analysis since those will affect the financial performance of a company.

With Kalmar Mobile Equipment, there should be correlation between low interest rate of customers' bonds and the sales volume. Relationship between interest rates and equipment sales could be due to the cost of capital. If one recalls the NPV calculation (equation 1) from the beginning of the chapter, then this would make sense. When interest rates are high or a company is not credit worthy, the company will have fewer possibilities with debt financing and must use more equity to fund its investments. Weighted average cost of capital (WACC) could be used in the NPV (equation 1) as r, for the required rate of return for investment, calculations when investments are leveraged using both equity and debt. WACC (equation 2) is the demanded return for an investment when the investment is financed with combination of equity and debt.

$$WACC = \frac{E}{D+E} * R_E + \frac{D}{D+E} * R_D * (1 - R_T), \tag{2}$$

where E is the sum of equity, D is the sum of debt, R_E is the required return for equity, R_D is the required return for debt, i.e., the interest rate of borrowed money and R_T is the corporate tax rate. Now one can see fast and clearly that when the proportion of E increases keeping R_E stable, ceteris paribus, will the WACC increase as well. If all parameters change, then determining the direction of movement in WACC will be harder without calculations. Normally it has been thought that R_E will be higher when the proportion of debt is high. One thing to note is that in all equity firms R_E is not particularly small either. (E.g. Brealey 2014, 221-226.)

Interest rate is not only factor to explain varying sales volumes. Seasonal effects are one to mention, as are foreign exchange rates, global container traffic and so on. One must also include company specific factors for example incapability to deliver machines at one period and in such cases the delivery is postponed to another period. Based on accounting principles, the revenue is recognized when the machine is delivered. The customer base might as well have explanatory power since need for new machines could be due to increasing customers' businesses. Industry specific factors should explain performance of sales of certain machines. For example, a need for heavy-duty forklift trucks could be due to increase in steel consumption and customers need to scale-up the production.

2.3 Research context analysis

After the financial crises, there has been large volatility in the cargo handling business in terms of sales and container traffic itself. From container throughput data, it is possible to see the effect of financial crisis; the container traffic reached the 2008 level two years later (Institute of Shipping Economics and Logistics 2017). When there is no organic growth in the business, there is no need for other investments than replacement investments, i.e., no reason to purchase more equipment than is disabled (Cargotec 2010).

In the sales volumes, one can see the effect of financial crisis to cargo handling and especially in the case of mobile equipment. The initial investment of a mobile equipment machine is relatively small and easily postponed or advanced based on the company's and macroeconomic situation. In addition, the lead-time of mobile equipment investment can be relatively short. Time from the request to delivery might be quite short from the customer's point of view and due to this; customers can wait for the most appropriate time for them to invest. Short lead-times will make sales forecasting and resource planning more difficult. All of these will have a combined effect to the sales volume. (Cargotec 2010.)

Highly cyclical sales behavior can be explained with external, customer related factors, and internal ones. Sales volume could drop due to incapability to deliver machinery to customers. Root cause for this is usually the resource planning or the performance of the factory. In global business, high proportion of work is allocated to subcontractors. In such case, the actual machinery company does the final stage of the manufacturing process, the assembly. From that one can understand the challenge, if all the suddenly there is a demand shock for equipment, there will be no subcontractor recourses to fulfill the demand and those orders will be processed in the next period when the recourses can be increased and hence the sales will increase as well. The customer perspective to cyclical sales is that even if the machine company would want to deliver the equipment, maybe the customer is reluctant to have it now; maybe they do not have any usage for it now and they would prefer later delivery.

To conclude this idea, exceptionally high sales volume can be due to the incapability to deliver equipment to customers in history. There could be also taxation driven reasons behind the cyclical sales. In some EMEA countries, e.g., in Finland, Sweden and UK, if the customer has the delivery of the machine before the year-end, customer can deduct the full depreciation in taxation and minimize the income tax payed. On the other hand, in some EMEA countries, e.g., in Germany and France, the deprecation rate is linear and due to that, there is not a similar tax benefit. There might be other seasonal effect, e.g., need to secure high cost budget for the next year as well and due to this, more costs are generated. In addition, if sales are ramping up towards the year-end, one might analyze this to be caused by personal incentive and performance plans.

First, all parties who have their personal performance of their incentive plan tied to sales, profits or other accounting period specific measure, will have the incentive to maximize the results on that specific period. Secondly, if executives have their personal incentives tied to share price or the market value of the company they do have an incentive to maximize the profits, the market value and the good news of their company at the end of the year. One could ask, is that not then less revenue in the next period? That conclusion is correct but rational human should maximize the net present value of their personal income. Due to discount rates, x amount of bonus one year from now is less than x amount today. (Strotz 1955.)

One must remember the negative effect of this, if one eats a whole cake today, there will be nothing for tomorrow. It applies here as well; invoicing everything in period X1 will decrease sales in period X2. Why one should even care about this. It is important to notice at the beginning of the study that the model will not be perfect and will not capture this internal behavior. The fundamental fact based model is incapable to model accurately year-end closing nor the January, since there has been human fixing to the results while the actual market fundaments have not changed. Even though the possibility of internal fixes is already identified, it will be hard task to model human behavior with macroeconomic factors and should be a separate study due to that.

2.4 Defining the scope for the research

In order to the research to be successful, a specific scope must be established. There are two decisions to be made, what and where. First is the where part, i.e., which geographical area will be covered in the research. Secondly, to which product line to focus on. In current setup, Kalmar has three areas to which the globe is divided: Americas, Asia-Pacific and Europe, Middle East and Africa (EMEA).

Before diving to analyze three regions and their special characteristics, it would be beneficial to have a good overall view of sales and order book development. On overall level, one can see quite stable development of order book and linear trend would be easy to fit. One interesting result is the ratio of sales to order book value per period. This can be seen as the rotation speed of the order book. If the ratio 1:1, then there is as much sales as there is order book. This would imply that the company can fulfill its orders in fast pace. If the ratio is small, then it takes long time for the company to fulfill its orders.

Large order book will give buffer for company to have machinery to deliver in later periods but should affect the order intake as well. If a delivery time is an important factor to a customer, then a company with a high order book might not be the partner to fulfill requirements of the customer who would prefer short delivery times. Normally companies will first deliver the already ordered items and then start the process for the new customer. There has been better completion ration of orders received in Americas and in Asia-Pacific than in EMEA. This can be due to internal issues or purely that the customers do not demand machines to be delivered as soon as possible and the machines can be stored in a warehouse for a while. The distance between customers and factory can have an effect to the completion ratio. When customer in Americas or in EMEA orders a machine that is manufactured in the factory located in China, the transportation will increase the delivery time and lower the completion ratio. While machine is in the warehouse, it is still as an order in the order book.

To decide where to focus is a two horse race. Asia-Pacific is not the easy one to model since collecting data will be highly challenging because Asia-Pacific is not a single legislative area. The choice is between Americas and EMEA. Americas is highly interesting but the consolidation of the product portfolio is challenging, there is roughly speaking only one product offered from the Mobile Equipment, the terminal tractors. From above analysis and the fact that there is useful data available through Eurostat and European Central bank, the EMEA market is selected.

2.5 Analysis about Kalmar: geographical regions and product lines

Mobile equipment division is a part of Kalmar and Kalmar is part of Cargotec Plc together with Hiab and MacGregor. Mobile equipment division can be split to different product lines, each presenting each individual product type. Each product line contains multiple profit centers for each subgroup of equipment in each product line. Profit center is a part of a company that is responsible for its profit. By definition, profit center's responsibility is to take care of the profitability, which includes revenues from sales and corresponding costs of the goods sold. Profit is then the left over when one subtracts costs from revenues.

Kalmar and Mobile Equipment has eight profit centers and three profit center groups or product lines. Three profit center groups are counterbalanced container handlers, forklift trucks and terminal tractors. In these are the eight profit centers. In counterbalanced container handlers, there are empty container handlers and reach stackers. Forklifts are divided to three profit centers based on the amount of weight the machines can lift. Light forklift trucks are one group; medium forklift trucks are in the middle of range and at the end of lifting capacity are the heavy forklift trucks. Terminal tractors are divided to three categories: medium terminal tractors, heavy terminal tractors and emerging market terminal tractors.

Forklift trucks are usually used in industrial work to lift and transfer materials or finished goods from one location to another but also in special applications, such as roll-on/roll-off (RoRo) applications. Especially the heavy forklift trucks are used in heavy manufacturing and construction industries like in steel industry, where the materials are extremely heavy and so are the finished goods to be transported to customers. Other industrial products transported with forklift trucks are paper and pulp, wood, steel, concrete and offshore products. With this listing, one could easily see that raw material and finished product indices could be the ones to explain partially the sales of forklift trucks.

On the other hand, one must remember the continuing technological progress of equipment. The latest technology can be faster but more importantly can have significantly lower variable costs. According to Kalmar, the new electrical medium size forklift truck has approximately payback period of 2 years and is estimated to run with 50 % lower cost than the current alternatives (Kalmar 2017). In addition, electric machine is quieter, less polluting and will have more operating environments than the equivalent diesel truck will. If there is no exact index to model the development of equipment, could one model the technological improvement by using interest rates as an explanatory variable? Idea would be that interest rates would transform the marginal development of equipment to time series form. As learned already, low interest rate for a firm implies more possibilities to invest. With this, marginally better investments would be possible to execute. For example, if a firm has a perfectly running machine it can be profitable to change the machine to something that has lower running cost if the cost of capital is low. From this point of view, interest rate are definitely the ones to use in the analysis.

To conclude this analysis and to lead further, analyzing and explaining sales of different machines is a combination of direct and indirect factors. Direct factors would be the equipment itself and its benefits, the demand for extra machines or the replace old ones, development of global trade, traffic and industries where machines are operated.

Indirect factors would be other non-business related macroeconomic factors, interest and FX rates, development of consumer confidence and demand of products that are distributed indirectly with the machines. The focus of this thesis is to analyze how the macroeconomic factors explain the sales volume of Kalmar mobile equipment.

In EMEA, both counterbalanced container handlers and forklift trucks are highly selling products and have a high order intake as well. This is easy to derive as well with common sense; Europe is large market and producer for goods that are transported via seaports. Large and well-being economy will also require equipment for construction and manufacturing industries, i.e., the forklift trucks.

To select between the two is not easy but has to be made. Forklift trucks are selected for the research because there are more monthly data available of the business related factors to be used for econometric modelling than there is for counterbalanced container handlers. From financial perspective, there is a similar cyclical sales pattern in both, but the forklift trucks order book has increased over the time. Possible internal issues to deliver machines to customers can further swell the order book of forklift trucks. Q3/2017 results were highly promising and one can only expect high Q4 as well, since the machines are ready but have to be billed and delivered in order to be sold. In terms of stability, counterbalanced container handlers' order book is highly stable, but have not increased that much, which raises the question, how to find explaining factors. The order book development of forklift trucks is better since its increasing and easier to explain since there is an actual change happened.

In some quarters there has been familiar controversial trend in the sales and orders received of the forklift trucks but phenomena is weaker than in the case of counterbalanced container handlers. With cyclic sales and high order intake, the order book has increased quite significantly. The order book will always increase when the orders received line is above the sales line. In the case of forklift trucks, there has been 6 quarters with higher sales than order intake out of 27. This is quite interesting finding and points out possible issues with sales and more precisely said, with the deliveries. Constantly higher order intake than sales, will imply ever-increasing order book, i.e., customers who are waiting their machines to be delivered and this will transform to longer lead-times for new orders. Especially in the case of year 2017, in Q3 sales were marginally higher than orders received. One can then easily observe that the order book has reached its all-time high as with the counterbalanced container handlers. The next step should be to realize the potential of the order book and improve the sales volume for the rest of the year and for the next year as well. In October 2017, there were more sales than new orders received, i.e., the order book will decrease as a whole. From experience, one could assume to see later a hockey stick effect Chen (2000, 186) to take place during Q4. The hockey stick effect refers to ramping year-end sales as a high hockey stick while playing ice hockey (Chen 2000, 186).

Forklift are mostly used in industrial and trade environments and the with-in those industries economic cycles can have serious affects to sales. First two years there were steady increase and then a large drop. After disappointing year 2013, the sales volume has had serious cyclicality and seasonal trend. Production of construction industry was in heavy turbulence since the late 2012 onward. In the forklift sales, one can easily see the hockey stick effect Chen (2000, 186) in the time series. Other interesting finding from the time series is the every other quarter peak. Since Q4/2014, there has been higher sales in every other quarter and Q3/2017 should have been such based on poor performance of Q2/2017. This decline in sales can be due to incapability to deliver equipment to customers or the fact that there have not been enough orders to fulfill. Latter can be found from order book and first from business operations. Determination of bottlenecks and issues in the delivery could be in this case be identified with the time the equipment spent in the warehouse. When one calculates a standard time for a machine to be in stock, it would greatly help to identify machines that have been in stock longer than machines typically are. Comparing machines to the normal time spent in warehouse would indicated whether there are machines that should already have been delivered.

When there is incapability to deliver machines during the year it will results large increase in Q4 sales because company will want to show the revenue on that accounting period. Since, if there are the normal sales amount plus the last period’s non-delivered machines, the combined result of those two for that particular period should be something to observe. Of course, the incapability to deliver equipment to customers can continue as well during the Q4.

After analyzing the internal factors, can the focus of analysis be in the actual sales, why was the sales low compared to previous periods and expectations. Finding the macroeconomic factors to explain sales volume is the purpose of this thesis and is useful for forecasting.

2.6 Next steps in the study

Idea behind the thesis is to model whether specific macroeconomic factors have an effect to Kalmar Mobile equipment sales. The idea and assumption is that the sales have causal affects from the macroeconomy and this assumption will be tested with one product line. When one can find causal relationship between products and the economy, the research is successful. In the future one could then extend the research to cover all profit centers and regions to have full coverage of mobile equipment division. Suitable models in such framework could be Vector Autoregressive (VAR) models and a long-run relationship model, a process introduced by Engle and Granger (1987).

Engle and Granger (1987) process assumes that different time series can be cointegrated and move together over the time. In the process, Engle and Granger (1987) argue that if the time series are cointegrated then the series do not separate for long periods but will seek each other after a shock. Cointegrated process can be extended to include multiple explanatory variables. In the context of this thesis that would mean using both financial factors as well as industry specific factors when for example explaining the sales of forklift trucks. Building cointegrated model will take time but if one finds the relationship then maintaining the model should be easy and performing forecast can be done without extensive econometric competence. An economist should do the maintenance and improvements of the model in order to have statistically sound model.

The focus of the thesis about the forklift trucks is in the heavy industry, international trade and the development of those markets and especially from the financial perspective. With-in the research both financial and business related perspectives are investigated and promising leads are reviewed. Financial factors could be the investment cycle of machines, FX rates that the customer companies are facing and interest rates of their debt. Business related factors are customer industries where the forklift trucks are operated, e.g., heavy industry, construction and manufacturing, materials and goods usage and movement. Goods movement is one important factor to investigate since goods and finished products are both lifted with forklift trucks. If one can derive a development index of forklift trucks that would be one explanatory variable as well. To derive a forklift performance indicator can be challenging task but at least in the future, that should be tried.

When analyzing order book and sales together, one can find relationships between orders coming in and sales revenue. There are always certain lead-times in manufacturing of industrial goods. With order book, it is easier to find those lags to test relationship between sales and macroeconomic factors. A rule of thumb seems to be that it takes roughly one or two quarters to transfer financials from the order book to income statement depending on profit center group. This means that peak in the order book in period Q1 will show as a peak in sales in period Q3 and in some cases in period Q2. In addition, the order book will be a powerful tool when understanding low sales periods. If sales are low but the order book is increasing, this will give an indication of the incapability to deliver equipment to customers. Complexity of manufactured equipment will affect the transfer time from the order book to sales. The order book do not imply the true lag length to use but will give the minimum where to start.

Even though there is a stable overall level development in sales and in the order book, the story is quite different when looking regions individually. Cyclicality in sales is usually unwanted effect of cyclical order intake and delivery incapability. If there is no peak in sales when the order book plummets, it means that there were fewer orders received than machines were delivered. When the order book plummets and the order intake is modest, it will mean hard times for next periods to come. Sales can only increase over the time when there is increasing trend in the order intake. This implies that in some cases, the order book will increase and sometimes the company will use its order book to compensate modest order intake. With the cyclical and peaking sales, it could be hard to find good factors to explain the sales performance and causal effects between sales and macroeconomy. Causal effects instead or pure correlation or association, between two variables is often the main interest of macroeconomist and other researches (Angrist, Imbens & Rubin 1996, 444). The issue of cyclical sales could be solved by implementing instrumental variables. If Y (sales) is the dependable variable and variable X (macroeconomic variable) cannot statistically acceptably explain Y but explanatory variable Z (order book or order intake) has statistically significant explanatory power to X but not for Y. One could formulate a model where Y is model with Z but through X. Then one could progress with the model forward and in the case of a forecast model, create a stable forecast for next periods and then alter the result based on historical behavior.

For example, if one uses Engle-Granger (1987) process and has a stable sales forecast for next year, it would be highly useful to manipulate the sales to behave in similar fashion as before. In the figure 2 below the idea is presented. If the forecast model generates stable forecast curve, the blue line in the figure 2, it would be highly useful to modify the forecast model to have similar cyclicality in forecast as there are in the actual sales. If there is pattern that for Q1 the forecast model overestimates the sales constantly and underestimates sales for Q4, then one could easily calculate deviations from the history and then calculate new so called reality corrected values for the forecast. In this case, the forecast would show what would be the sales volume if the sales would done without internal issues and machines are delivered quite evenly during the year. In addition to the fundamental based model, one would have the more realistic manipulated forecast model.

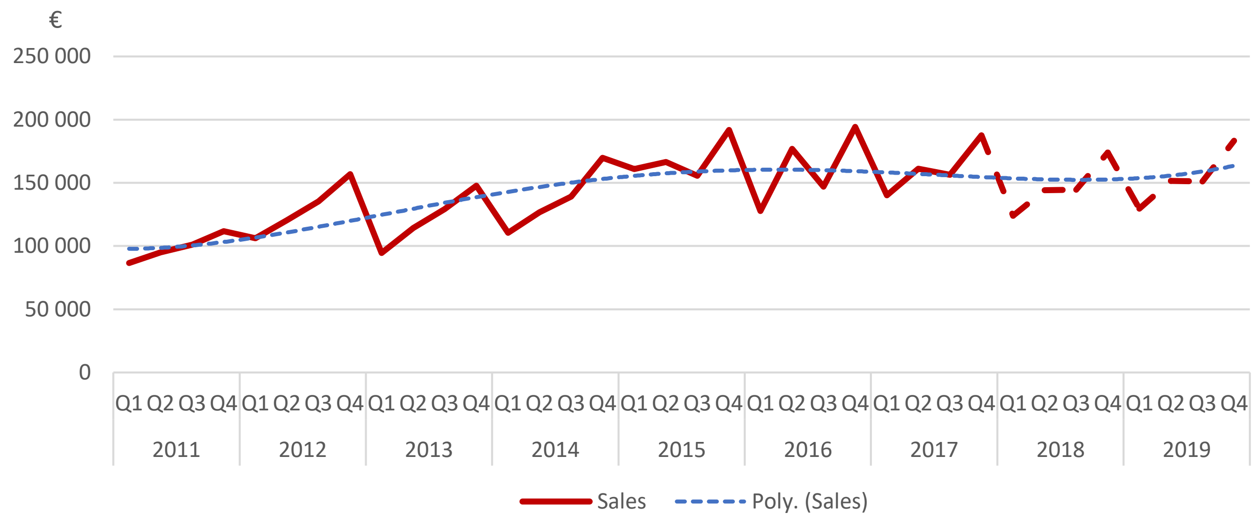


Figure 2. Hypothetic forecast model for a simulated sales series (estimated by the author).

In the figure 2, there are illustrated the results of the hypothetic forecast model together with simulated sales. Sales are estimated from historical sales data by using a mix of sales entities and their performances. The estimated forecast model is a trend function to the power of six to replicate the steady forecast. The forecast error between actual sales and forecast can be observed from figure 3. When the actual econometric causal effect model is quite stable without too much cyclicality, then the cyclicality must be brought from somewhere else. In this case, the deviation between the estimated sales and the actual sales is calculated and illustrated in figure 3.

From the figure 3, one can see that the model overestimates the sales for Q1 for most of the time and underestimates the sales for Q4. This can be seen since the column values are below 100 % for Q4 and correspondingly above 100 % for the Q1. Value is calculated based on forecast value / actual value. However, for Q2 and Q3 the forecast model do perform accurately. When the forecast model provides too smooth forecast, one could manipulate the forecast results to present estimation that is more realistic. Since, when the forecast error is systematic, it can be corrected. Of course, if the company changes its operations, this manipulation would not work anymore and company would be better off by sticking with the original forecast model.

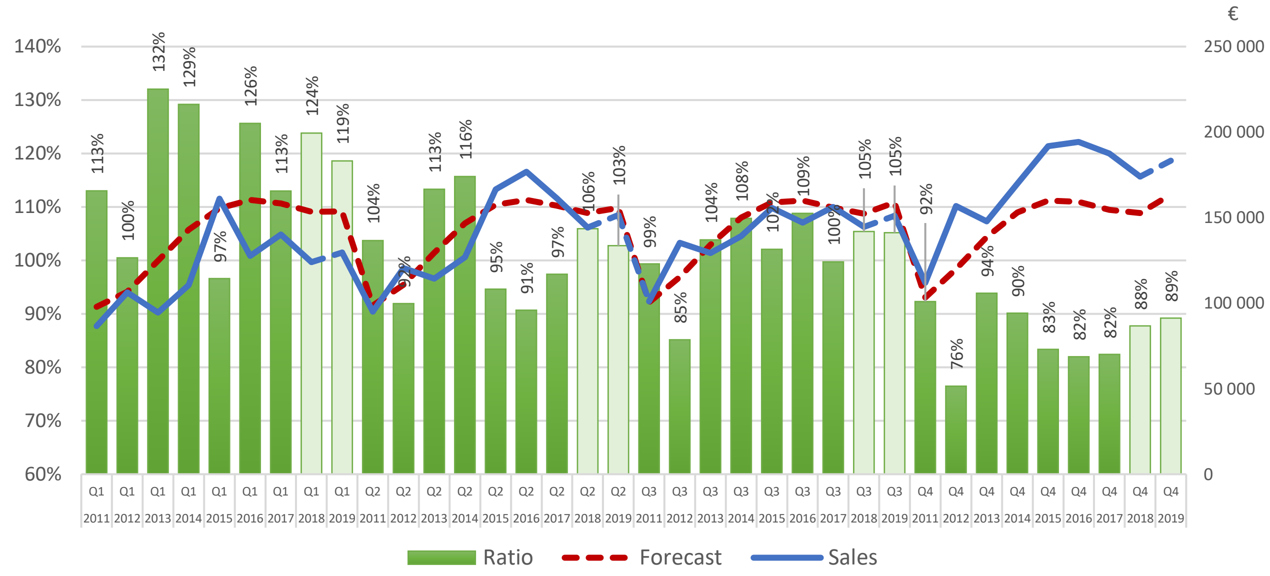


Figure 3. Forecast errors grouped to quarters (estimated by the author).

Manipulating forecast with historical patterns could be one method or one could derive a true demand for a company (Gilliland, Sglavo, Tashman 2015, 82-86). The true demand approach is highly interesting and could be useful in these cases. In addition to the more simple calculation methods presented by Gilliland et al., (2015), Chockalingam (2009) presented two approaches to calculate the true demand. One starts with observed bookings, i.e., from an order book and other one starts the calculation with observed (gross) shipments. Methods are:

$$
\begin{aligned}
&1)\ \boldsymbol{Observed\ Bookings}\\
&-Requested\ deliveries\ in\ the\ future\\
&-Exaggerated\ customer\ orders\\
&=True\ demand
\end{aligned}
$$

$$
\begin{aligned}
&2)\ \boldsymbol{Observed\ (Gross\ Shipments)}\\
&+Cuts\ (unfilled\ orders\ that\ are\ cancelled)\\
&+Backorders\\
&-Carryovers\\
&=True\ Demand
\end{aligned}
$$

Idea of the true demand approach is fascinating but whether it is compatible with a machinery company is something to thought and analyze next.

The first of two Chockalingam (2009) definitions has the base from the order book, in the formula the observed bookings. Next, there is elimination for deliveries to happened in the future. When these two factors are summed, one will have the wanted deliveries to happen in this period. After this, there is the elimination of exaggerated customer orders that are not expected to happen at all. This approach is compelling but for the first method, one will need the time series for the future deliveries. When the goods are manufacturing items, determining the delivery accurate might be challenging and to have time series data of that is challenging as well.

The second approach derives the true demand from another point of view. The second approach has base from gross shipments, i.e., from the sales. To this is then added amount of orders that would have been delivered now but the customer have cancelled it. Then backlog orders from other periods are added, i.e., the orders that were supposed to be already be delivered but are not. To finalize the calculation, the carryover orders are decreased because customer has request to change the delivery.

The true demand approach is highly compelling and brings some sense to the sales behavior. With this, one could manipulate the sales data to have something more fruitful to model with econometric models.

There are many forecasts and forecasting periods. In addition to the forecast horizon, a forecast accuracy is important. The accuracy of the forecast should be relative and not about absolute values; one million error in a monthly forecast is different from one million error on a yearly level. From forecast accuracy to forecast periods, there are two or three different types of forecasts identified. First are short- and medium-term forecasts. With-in the scope of the research is the beginning of the medium-term forecast, ideally around one year forecasting capability. Second is a long-term strategic forecast, which will give an answer to a question about direction the industry is progressing and how we should relate to it. Where the medium-term forecast would be created on a monthly basis, the long-term forecast could be created or at least aggregated to a yearly level. Ideally, the long-term forecast would have a view for next five years.

To recap and conclude some thoughts so far. The research will focus on forklift trucks in geographical area consisting Europe, Middle East and Africa. In the study, especially heavy industry and business facts are reviewed. If there were problems to model the sales as it is, it would be possible to create a model using instrumental variables that can be modeled with macroeconomic variables. One possible issue for econometric modelling would be cyclicality of sales. In such case, one could manipulate the stable macroeconomic forecast to represent the historical cyclic behavior around the forecast model. Third alternative would be the true demand approach to manipulate sales time series before econometric modelling (Chockalingam 2009; Gilliland et al., 2015).

3 Econometric modelling

3.1 Forecasting practice

Whether to use explicit judgement for forecasting has been a topic in research since 1970s (Bunn & Wright 1991, 501). The judgement is highly fascinating since there should be at least some sort of judgement to be made but what is correct and what is not, is the real question. One of first studies in the field to find the best practice for sales forecasting was Rothe (1978). In the study, Rothe (1978) found out that 50 out of 52 interviewed companies used judgmental forecasting models or methods in some extent for forecasting.

Next step in the research was extensive survey study by Klein and Linneman (1984). In their study, Klein and Linneman (1984) interviewed 500 of the world's largest companies to understand their forecasting practices and the caveats experienced during forecasting. Klein and Linneman (1984) found out that companies had experienced large difficulties and caveats when using only statistical models. Cerullo and Avila (1975) found similar result earlier in their Fortune 500 research. Cerullo and Avila (1975) took a draw from Fortune 500 list and had 110 companies for their survey. Their key finding was that 89 % used judgement exclusively or combined with another sort of forecasting model (Cerullo & Avila 1975).

From the previously mentioned studies, one should not take too far-reaching conclusions, econometrics and forecasting has evolved since the studies were made. One thing to note is that the current management of companies have been studying at the university with the knowledge and information available during these studies. Whether the management has studied further the forecasting practice, might explain at least to some extent the lack of statistical methods in business forecasting. Management could be skeptical for new methods that younger employees bring to the company and be reluctant to have forecasts created using those methods.

One key point hidden in the previous paragraphs is the actual level of judgement and the object what is going to be influenced. Based on analysis by McNees and Perna (1981), Corker, Holly and Ellis (1986) and Turner (1990), one will normally observe the human judgement for a model specification error or to model a structural change that the model did not capture (Bunn & Wright 1991, 502). Related to the model specification, Reinmuth and Guerts (1972) found in their study that unconventional events will be better forecasted and with higher accuracy when judgement is applied. One should then question, should the unconventional events be modelled and not just adjusted based on experience. Reinmuth and Guerts (1972) found that for example sales promotions and sales forecasts would benefit from judgement-based adjustments. Experts in their respective field will increase the forecast accuracy of the sales forecast when one will imply their expertise by the judgmental adjustment. This would imply that the so-called best practice for forecasting would be a statistical model combined with human adjustment, naturally to both directions in the case of sales. (Bunn & Wright 1991, 503.)

Normally and in the context of this thesis, judgment and adjustment are related to manipulating the outcome of the forecast, the actual reported numbers. However, later in this thesis there are many steps, which can, and will be read as judgements. What variables and model to use, what kind of model to use and so on. All choices and selections can be viewed as judgmental adjustments, e.g., Dawes (1975), Armstrong (1985), Bunn and Wright (1991). Bunn and Wright (1991) identified two other judgmental areas, which have positive human interaction; those are parameter estimation of the econometric model and the data analysis.

Data selection and model creation should be the judgmental process and not that much the manipulation of the sales forecast, since who would prefer working towards to create a model and then someone else would manipulate it to a direction what is wanted to see. Possibly even worse scenario is when the forecaster will manipulate the results, in that context there is then something wrong with the model and the model should be further specified if possible.

Besides the split between model specification error and structural change, one can split the forecasts to objective and subjective where the latter means the judgmental forecasting which is created with experience and the previous mentioned is the statistical method to forecast (Webby & O'Connor 1996, 92). Webby and O'Connor (1996) reported from extensive literature review that 40 to 50 percent of forecasts in time series forecasting is done with subjective forecasting techniques. Humans do have great capabilities to understand patterns and to find cause-effect relationships between variables. Humans do have good capabilities for trend recognition and to search causality and the usage of those can improve the forecast accuracy compared to pure objective forecast. In addition, one should exploit the high capabilities of humans to model and understand the discontinuities from the past in the time series. (Webby & O'Connor 1996, 93-98.)

Based on research Turner (1990), Donihue (1993) there are human interaction and judgment implemented to the forecasts. Different judgmental changes are made to incorporate information outside the model specifications. Objective for these adjustments is to have better forecasting accuracy but interestingly the adjustment is done for the model and not for the output. Interferences are both frequent and successful.

Interference to the model can be split in the case of objective forecast, i.e., with the statistical forecast model to three categories. These are non-contextual adjustment, contextual adjustment and structured adjustment. The non-contextual adjustment is the unwanted effect for the forecast from the point of view of this thesis since the non-contextual adjustment is not fact nor fundament based adjustment, but more of a hunch. One might argue on the behalf of non-contextual adjustment but why would one adjust the fact-based model based on one's intuition.

The objective for this thesis is to create a fact-based model to eliminate the intuition based forecasting and only use facts. Contextual adjustment is done when there are extra information outside the model available and the forecaster can rely on one's expertise to adjust the model to have higher forecast accuracy. In such cases, Mathews and Diamantopoulos (1986) report the judgmental adjustment to be effective. With the structured adjustment, one is adjusting the forecast with external information but one person creates the forecast and another person does the adjustment. The process will not always improve accuracy and Bunn and Wright (1991) criticized its ad-hoc nature. (Webby & O'Connor 1996, 103-104.)

3.2 Cointegration

Interaction and cause-effect relationship between two variables is the main idea for this study. When there is a causal relationship between two variables, will the one reveal and indicate to movement of the other variable. (Gourieroux & Jasiak 2001, 95.) One can search and model the causality with econometric analysis.

For this study the methodology to find causal effect between different variables, is the Engle-Granger cointegration. Engle and Granger (1987) presented their development for cointegration. A year earlier Granger (1986) published a paper where he argued on the behalf of the theory of cointegration, which Granger had presented for the first time in 1981. According to Granger (1986), it will make sense that some variables are cointegrated, and when those truly are; those should not separate too far from each other for long periods, at least not on average, i.e., in the long-run. With that statement, there is an important factor to point out; even the cointegrated variables will drift apart from each other in the short-run but not over the time. One must already make this note, this effect will be there when one performs forecast for the financial series of interest, e.g., for sales. This drifting can cause R^2 to be lower than 100% and closer to 50%. The goodness of the forecast model is dependable, how well one can model the history and how well do the cointegrated variables explain and fit to each other.

Normally economics theory will point the pairs or groups of variables that fulfill the requirements of cointegration (Engle & Granger 1987, 251). In the context of this paper, there might not be that much specific theory to find the variables but analyst can find appropriate variables to start the study. For the forklifts trucks for example, natural explanatory variables are indices that represent the geographical area of study and the customers' businesses. With forklift trucks, one will lift heavy materials in manufacturing and construction industries and for example in mining industry. Naturally, forklifts are used in trade goods business, goods must be moved from one plane to another, loaded to trucks and moved with-in warehouses and distribution centers.

All variables in this study are time series variables, which usually are not stationary but non-stationary series (Maddala & Kim 1998, 20). When variable is not stationary, it is integrated with an order of d, $I(d)$. The power of integration is the count of differences needed to take in order to have stationary series. If variable is $I(1)$, then one difference will make the series stationary, i.e., d time difference is needed in order to have stationary series of variable, which is $I(d)$. (Maddala & Kim 1998, 25.) When series is for example $I(2)$, integrated order of two, then one must have two unit roots in the time series of that variable (Verbeek 2008, 282).

If one can find a linear relationship between $y(t)$ and $x(t)$ when both are $I(1)$ and the linear combination is stationary, i.e., $I(0)$ the residual $u(t)$ is then the realization of that linear combination found. When $u(t)$ is $I(0)$, the $I(1)$ variables $y(t)$ and $x(t)$ are cointegrated. If the $u(t)$ is $I(1)$, then the residual has a unit-root and the model has spurious regression. (Maddala & Kim 1998, 21.) Indication of spurious regression, a model that do not have actual meaningful usage, can be firstly be identified with a high R^2 figure and secondly by a low Durbin-Watson statistics. High R^2 would imply that the model fits the data well but the low Durbin-Watson implies that there is a large amount of positive serial correlation (Maddala & Kim 1998, 28). Positive serial correlation means that a positive deviation is followed by another positive deviation in the residual.

Testing for cointegration can be done with different ways (Maddala & Kim 1998, 28). Engle and Granger presented (1987) test procedure for cointegration, which is residual oriented. In this test, one estimates a model between presumably cointegrated $I(1)$ variables and saves the residual. One then performs a unit-root test for the residual to test, whether there is real cointegration between variables in the model. The null hypothesis for the test is that there is a unit-root and alternative hypothesis is that there is no unit-root. Critical values used in this test developed by Engle and Granger (1987) are specially computed for the purpose.

To test whether a variable is stationary, i.e., does it have a unit-root; Dickey-Fuller test is presented. Dickey and Fuller (1979) have developed a unit-root test that uses variables itself and performs a simple regression to test its stationarity. Regression model is run with or without a constant and a trend. Null hypothesis for the regression and test is that there is a unit-root. Alternative hypothesis is that there is no unit-root and series is stationary. The Dickey-Fuller test is performed to test, whether the ρ in (equation 3) deviates from one or not.

$$y_t = \rho y_{t-1} + e_t, \tag{3}$$

where $y(t)$ is the value of the variable at moment t, ρ is the coefficient, which is tested, for the lagged value of the variable $y(t)$ and $e(t)$ is the error term in regression model. The null hypothesis in this sense is that $\rho = 1$ and the alternative hypothesis is that, it is not. The Dickey-Fuller test is derived from (equation 3) by adding $-y_{t-1}$ on both sides of equation, which will result the (equation 4). With arranging the coefficients, one will have from (equation 4) first (equation 5) and ultimately (equation 6), which is the Dickey-Fuller test equation and an OLS regression equation as well. One will then perform an OLS estimation for (equation 6),

$$y_t - y_{t-1} = \rho y_{t-1} + e_t - y_{t-1} \tag{4}$$

$$\Delta y_t = (\rho - 1) y_{t-1} + e_t \tag{5}$$

$$\Delta y_t = \delta y_{t-1} + e_t \tag{6}$$

(Dickey & Fuller 1979.)

When one extends the Dickey-Fuller test with more lagged values of the dependable variable, it will lead to a test model called augmented Dickey-Fuller test (equation 7). Test formula for example when testing an AR(2) model is:

$$Y_t = \delta + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \varepsilon. \tag{7}$$

Test formula (equation 7) can then be presented in the following way (equation 8):

$$(1 - \phi_1 L)(1 - \phi_2 L)(Y_t - \mu) = \varepsilon_t \tag{8}$$

If the variable is stationary, the coefficients ϕ_1, ϕ_2 must both be in absolute terms less than one. If one of coefficients is equal to one then one has one unit root in the variable, if two coefficients are equal to one then one has two unit roots. To test, whether there actually is a unit-root in the variable, one can do that by using OLS estimation. When the original augmented Dickey-Fuller formula is presented in the following way (equation 9), one has the OLS regression for stationary testing,

$$\Delta Y_t = \delta + (\theta_1 + \theta_2 - 1)Y_{t-1} - \theta_2 \Delta Y_{t-1} + \varepsilon_t. \tag{9}$$

The augmented Dickey-Fuller test is the following, one will test whether coefficient of $(\theta_1 + \theta_2 - 1)Y_{t-1}$, i.e., $(\theta_1 + \theta_2 - 1)$ differs statistically from zero. Hypothesis for testing are the following:

H0: There is a unit-root in the sample

H1: There is no unit-root in the sample.

In terms of test statistic, this can be shown that when (equation 10),

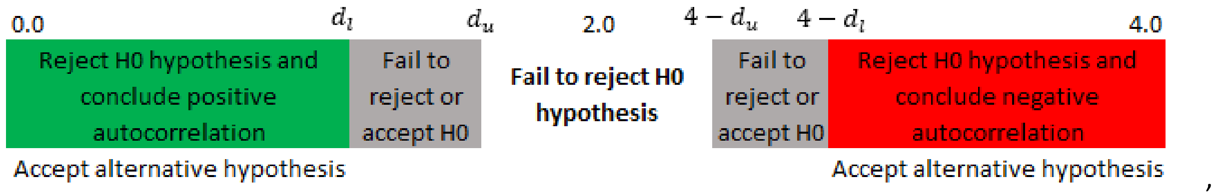
$$\pi \equiv \theta_1 + \theta_2 - 1 = 0, \tag{10}$$

there is a unit-root in the sample and if the previous (equation 10) do not hold, there is no unit-root in the sample. The main idea of additional lags in the augmented Dickey-Fuller test is to have an error term, which is white noise process asymptotically. The white noise process is a requirement that the distributional results and conclusions are valid. As usual, one should not include to model, in this case to the ADF regression model, any more variables than are necessary. Additional lags do lower the power of the test process and if possible, one could test the test model with Akaike's Information Criterion. (Verbeek 2008, 286-287.)

To test the usage of variables, one must test also the autoregressive nature of the variable. Durbin-Watson statistics can be used for this test. One will derive the Durbin-Watson test results from the very same regression that was used for the unit-root testing. Durbin and Watson presented their test procedure for autocorrelation (1950, 1951, 1971), which uses the residuals from the autoregressive (equation 9) regression. Durbin-Watson test statistic d is (equation 11)

$$d = \frac{\sum_t^T (e_t - e_{t-1})^2}{\sum_1^T {e_t}^2} = 2(1-r) - \frac{{e_1}^2 + {e_T}^2}{\sum_1^T {e_t}^2}, \tag{11}$$

where r denotes the first-order autocorrelation in the variable. When one has reasonably large sample, then the later part of (equation 11) will be marginal and the (equation 11) will converge to $2(1-r)$ (Greene 2012, 963). Durbin-Watson statistic is not without faults, it will have two areas where test will not accept or reject the H0 hypothesis. Durbin-Watson statistics is a scale from zero to four but the ideal value is two, which indicates the absence of autocorrelation in the variable. Scale of statistic is presented in the figure 4:



where d_l denotes the lower bound and d_u the upper bound of the area of uncertainty.

Figure 4. Interpreting the Durbin-Watson test static (compiled by the author).

For example, when one has ~ sample size of 80 and 4 variables, the upper bound for Durbin-Watson statistics is ~1.7.

3.3 Modelling

To use variables in a regression model, one must first test whether the variables are stationary and whether there is autocorrelation with-in the variable. From previous paragraphs, one knows that these are tested with Dickey-Fuller and Durbin-Watson tests. Testing for stationarity with Dickey-Fuller test the augmented version is applied.

Table 1. Augmented Dickey-Fuller test for Calendar ManufacturingUpdate; regression of DCalendar ManufacturingUpdate on:

	Coefficient	Std.Error	t-value
Calendar ManufacturingUpdate_1	-1,3579	0,1736	-7,8236
Constant	130,0300	16,6350	7,8171
Trend	0,1964	0,0333	5,8922
DCalendar ManufacturingUpdate_1	0,4185	0,1304	3,2083
DCalendar ManufacturingUpdate_2	0,1559	0,0992	1,5721

	5 %	1 %	
Critical values used in ADF test:	-3,453	-4,048	
ADF-Calendar ManufacturingUpdate	-7,824	**	
DW critical value	1,78	H0	H1
DW	1,9	Accept	Reject

Count of variables	5		
Count of observations	101		

Source: Estimated by the author.

From table 1 one can see that variable manufacturing is stationary, the ADF-test result is -7.824, which is smaller than the 1 % critical value. Hypothesis for the ADF-test are:

H0: There is a unit-root in the sample

H1: There is not a unit-root in the sample

Since the test results is lower than the critical value, one must reject the null hypothesis and conclude that there is no unit-root in the sample and the variable is stationary. Durbin-Watson statistic was for testing autocorrelation in the sample. Hypothesis for the Durbin-Watson test are:

H0: Series values are not correlated with each other

H1: Series values are correlated with each other

With the Durbin-Watson test, one must remember that test value that fails to reject null hypothesis, is a value between upper bound or four minus upper bound and two. Since test statistic is 1.9, which indeed is between the critical value 1.78 and 2, one must conclude that the test fails to reject the null hypothesis. Since null hypothesis is not rejected, one can conclude that series values are not correlated with each other.

Table 2. Augmented Dickey-Fuller test for Calendar ConstructionUpdate; regression of DCalendar ConstructionUpdate on:

	Coefficient	Std.Error	t-value
Calendar ConstructionUpdate_1	-0,73061	0,13458	-5,4289
Constant	71,986	13,485	5,3384
Trend	-0,021511	0,027285	-0,78835
DCalendar ConstructionUpdate_1	0,3013	0,10797	2,7907
DCalendar ConstructionUpdate_2	-0,072707	0,10223	-0,71118

	5 %	1 %	
Critical values used in ADF test:	-3,453	-4,049	
ADF-Calendar ManufacturingUpdate	-5,429	**	

DW critical value	1,78	H0	H1
DW	1,972	Accept	Reject

Count of variables	5		
Count of observations	101		

Source: Estimated by the author.

From table 2 one can see that variable construction is stationary, the ADF-test result is -5.429, which is smaller than the 1 % critical value. Hypothesis for the ADF-test are:

H0: There is a unit-root in the sample

H1: There is not a unit-root in the sample

Since the test results is lower than the critical value, one must reject the null hypothesis and, concluded that there is no unit-root in the sample and the variable is stationary. Durbin-Watson statistic was for testing autocorrelation in the sample. Hypothesis for the Durbin-Watson test are:

H0: Series values are not correlated with each other

H1: Series values are correlated with each other

With the Durbin-Watson test, one must remember that test value that fails to reject null hypothesis, is a value between upper bound or four minus upper bound and two. Since test statistic is 1.972, which indeed is between the critical value 1.78 and 2, one must conclude that the test fails to reject the null hypothesis. Since null hypothesis is not rejected, one can conclude that series values are not correlated with each other.

Table 3. Augmented Dickey-Fuller test for LEU28 MOVEMENT QUANTITY_IN_100KG/1000; regression of DLEU28 MOVEMENT QUANTITY_IN_100KG/1000 on:

	Coefficient	Std.Error	t-value
LEU28 MOVEMENT QUANTITY_IN_100KG/1000_1	-0,68963	0,16513	-4,1763
Constant	9,9276	2,3743	4,1813
Trend	0,0008445	0,0002516	3,3565
DLEU28 MOVEMENT QUANTITY_IN_100KG/1000_1	-0,5103	0,1547	-3,2987
DLEU28 MOVEMENT QUANTITY_IN_100KG/1000_2	-0,3851	0,13713	-2,8083
DLEU28 MOVEMENT QUANTITY_IN_100KG/1000_3	-0,10912	0,094923	-1,1495

	5 %	1 %	
Critical values used in ADF test:	-3,454	-4,051	
ADF-Calendar ManufacturingUpdate	-4,176	**	

DW critical value	2,2	H0	H1
DW	2,048	Accept	Reject

Count of variables	6		
Count of observations	101		

Source: Estimated by the author.

From table 3 one can see that natural logarithm transformed variable movement is stationary; the ADF-test result is -4.176, which is smaller than the 1 % critical value. Movement is calculated as a sum of imports and exports. Hypothesis for the ADF-test are:

H0: There is a unit-root in the sample

H1: There is not a unit-root in the sample

Since the test results is lower than the critical value, one must reject the null hypothesis and, concluded that there is no unit-root in the sample and the variable is stationary. Durbin-Watson statistic was for testing autocorrelation in the sample. Hypothesis for the Durbin-Watson test are:

H0: Series values are not correlated with each other

H1: Series values are correlated with each other

With the Durbin-Watson test, one must remember that test value that fails to reject null hypothesis, is a value between upper bound or four minus upper bound and two. Since test statistic is 2.048, which indeed is between the critical value 2.2 and 2, one must conclude that the test fails to reject the null hypothesis. Since null hypothesis is not rejected, one can conclude that series values are not correlated with each other.

Table 4. Augmented Dickey-Fuller test for LRM01Ev2; regression of DLRM01Ev2 on:

	Coefficient	Std.Error	t-value
LRM01Ev2_1	-1,7455	0,30755	-5,6754
Constant	15,206	2,6779	5,6785
Trend	0,0082593	0,0021368	3,8653
DLRM01Ev2_1	0,66745	0,25309	2,6372
DLRM01Ev2_2	0,4961	0,21393	2,319
DLRM01Ev2_3	0,4045	0,16568	2,4415
DLRM01Ev2_4	0,088828	0,11693	0,75967

	5 %	1 %	
Critical values used in ADF test:	-3,467	-4,077	
ADF-LRM01Ev2	-5,675	**	

DW critical value	1,83	H0	H1
DW	1,932	Accept	Reject

Count of variables	7		
Count of observations	79		

Source: Estimated by the author.

From table 4 one can see that natural logarithm transformed sales variable is stationary, the ADF-test result is -5.675, which is smaller than the 1 % critical value. Hypothesis for the ADF-test are:

H0: There is a unit-root in the sample

H1: There is not a unit-root in the sample

Since the test results is lower than the critical value, one must reject the null hypothesis and, concluded that there is no unit-root in the sample and the variable is stationary. Durbin-Watson statistic was for testing autocorrelation in the sample. Hypothesis for the Durbin-Watson test are:

H0: Series values are not correlated with each other

H1: Series values are correlated with each other

With the Durbin-Watson test, one must remember that test value that fails to reject null hypothesis, is a value between upper bound or four minus upper bound and two. Since test statistic is 1.932, which indeed is between the critical value 1.83 and 2, one must conclude that the test fails to reject the null hypothesis. Since null hypothesis is not rejected, one can conclude that series values are not correlated with each other.

To conclude variable test statistics, all selected variables are based on descriptive statistics testing stationary and do not have autocorrelation. All explanatory variables have a sample size of 100, which is normally assumed to qualify as a large-sample. Only the dependable variable, LRM01E has a sample size of 80.

Sales forecasting model is created with external variables using lagged values of those. From previous analysis and using business insight combined with common sense, one will find selection of variables more or less obvious. For the model, such overall indices are selected that present the actual environment and the usage of forklift trucks.

First variable is a manufacturing index for the EU28 area. EU28 stands for the 28 member states that belong to European Union. The manufacturing index present all manufacturing activities with-in the EU28 area excluding manufacturing of electricity and mining of salt and so on. All other manufacturing categories: food, beverages, clothes, electrical equipment et cetera are included to this index. This variable is used as an index where year 2010 is the base line, level 100, and the series is calendar adjusted. Idea behind this variable is that using manufacturing index; one will have a good overall index to connect forklift trucks to the items, which are lifted with that forklift truck in the manufacturing part of goods movement.

For goods movement in overall level, there is natural logarithm transformed variable called Movement quantity in kilos. This variable is calculate based on the sum of imports and exports. In economics, many are often interested about the balance of imports and exports, whether country or area is a net exporter or importer. For this thesis, the overall goods movement is needed, the sum of imports and exports. Both are lifted at some point with a forklift truck, no matter whether those are consumed with-in the EU28 area or not. By taking natural logarithm, one will have a good usability of the variable, since then both variables, movement and sales are natural logarithms. When both variables, dependable and explanatory, will one percent point increase in explanatory variable, indicate, how many percent points the dependable variable will increase. The amount is indicated by the coefficient of the explanatory variable.

The third variable for the model is a construction index for the EU28 area. The construction index has same setup as the manufacturing index. The construction index combines as name states, construction activities with-in the EU28 area. Similarly as the manufacturing, the series is calendar but not seasonally adjusted to capture the true nature of activities in the construction sector. The construction index is highly usable for the model since in many construction activities one will lift materials and move from one place to another. In addition, even if not all construction site have a forklift truck, the forklift truck will be used in the supply chain. When one builds a block of flats from elements, the elements and the raw materials will, of which the element is first constructed, be lifted and transported with a forklift truck.

The already mentioned variables are the explanatory variables for the model but one must have a dependable variable as well. With the forklift trucks, the model is created to forecast external sales revenue in the EMEA area. Variable is presented with natural logarithm transformation of the external sales in euros. There is slight issue with using euro values for the variable since not all EU28 countries belong to EMU area and use euro, but the reason is sound; one cannot sum local currencies together to have one aggregate variables.

One could model quantities instead of monetary value but then one would not have all the information available. When one models monetary value of sales then all optional extras are in the model, which customer have purchased to improve the stock machine. Optional extras are on a basic principle similar features, which a consumer would purchase to his or her car but with industrial machinery, the options are a bit different. In addition, when one uses the monetary value, then the customers' willingness to pay will be in the model. If the sales increase heavily, then it can be because of more machines were sold in terms of quantity, more equipped machines were sold or less discounts were given.

For this study, one will need lagged values of the explanatory variables. Lagging is highly useful and mandatory since there are always lead-times for production. In rare occasions, there is a stock item to sell to the customer but in most cases, forklift trucks are built on demand and are customer specific at least to some extent. This tailoring will cause a natural delay between the initial order placement and the delivery of the machine. If one assumes that it will take four months to build a forklift truck, should one then use four months lagged values of explanatory variables. Not exactly. To lag only the length of production would imply that executive of a customer woke up and instantly decided to purchase a forklift truck. There will be always business plans, investment plans and timeframes for executive's plans and desires. From this point of view, one should look lagged values more than four months. During data analysis was discovered that there seems to be a two-quarter lag between cyclical orders received and corresponding peak in sales. From this point of view, the minimum lag to be used is six months.

The next phase is to estimate the model, which is the most challenging task in the research. One has three different explanatory variables and lag length possibilities from zero to multiple tens. There would be thousands or possibly over one million estimations to be done. To find the possible lag lengths for the model, one could use an automatic model selection to narrow down possible lags for each variable. With automatic model selection, the econometric package OxMetrics runs all the possible combinations of the lag lengths to find the best model. When this is run for each variable, possible lags for each explanatory variable can be significantly narrowed down. One must review the lag lengths and not blindly trust the software, since the results might not make any sense. When one knows that it will take four months to manufacture a machine, then under four lag lengths will not make any sense.

Automatic model selection indicated possible lag lengths for each variable but those must be then combined for a one complete model. In the next phase, all lags and few extra were added to the estimation to have different combinations. For example, if the lags for an explanatory variable were 11, 13 and 15, then all lags between 11 and 15 were added to the model and for all explanatory variables. With this large model, the automatic model selection were run again and the model had been estimated. The model consists explanatory variable manufacturing with two lags, 11 and 15, construction with lags 11 and 12 and the goods movement with lag length of 14. There is also a constant in the regression model.

In the figure 5, there are two elements of the model visualized. In the top part, there are the actual model and the actual sales time series together on blue and red colored lines. On blue is the econometric model estimated, fitted, and on red the actual time series of sales.

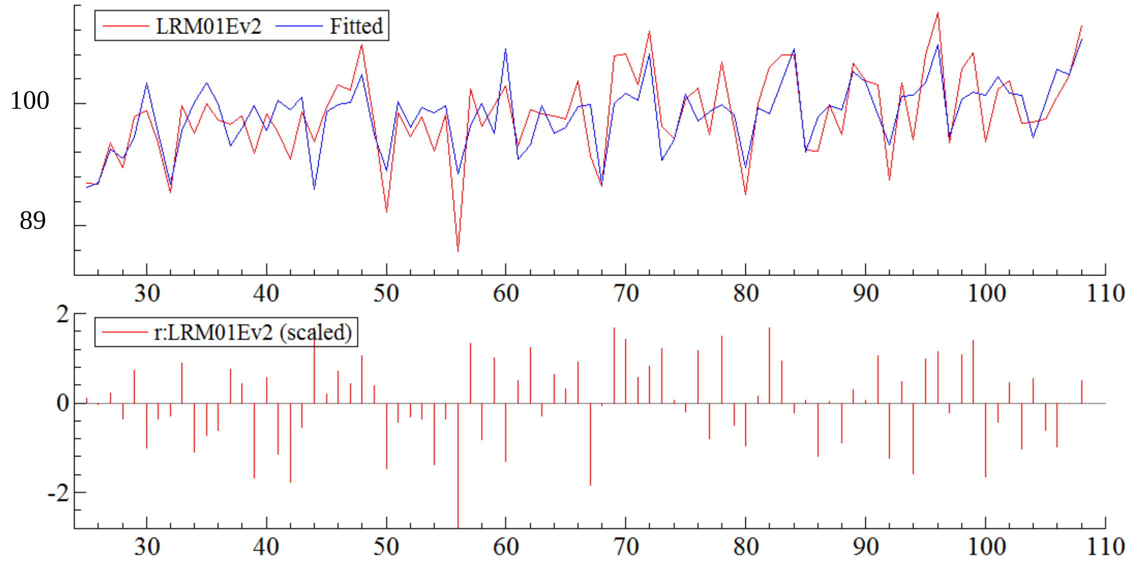


Figure 5. Forecast model (fitted) and the actual sales (LRM01Ev2) (estimated by the author).

From figure 5, one will see that there are deviations between the model and the actual series. There are few drops and peaks that the estimated model did not capture. Peaks and drops can, and will be, related to internal issues and timing of the sales recognition. The bottom part of the figure 5 is the residual analysis where the deviations between model and actual sales are visualized as scaled residuals. Visualization is useful to observe and understand the possible existence of autocorrelation. There are two types of autocorrelation. With positive autocorrelation, one will see same signed deviations, e.g., positive deviation will be followed by a positive deviation and vice versa. In the case of negative autocorrelation, negative deviation will be followed by a positive deviation. With the pattern in the figure 5 and statistical test introduced later on (equation 14), one can conclude the absence of autocorrelation in the residual.

3.4 Test results of the model

The test results of the estimated model are reported in the table 5. In the top part of table 5, are the variable estimations and tests. The variable estimations include regression coefficients for each lagged variable and for the constant as well. There are statistical significance testing for each explanatory variable on the same row as the explanatory variable is. Key statistics for the entire model are in the middle of table 5 and at the bottom are the residual tests. The middle part of table 5 includes goodness of fit measures and other measures of the model. Residual tests are particularly important in order to have statistically sound model for forecasting and to trust for the estimation done.

Table 5. Estimated forecast model – sales is the dependable variable.

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-33.6829	9.931	-3.39	0.0011	0.1285
Calendar ManufacturingUpdate_11	0.0300708	0.007533	3.99	0.0001	0.1697
Calendar ManufacturingUpdate_15	-0.0130077	0.004827	-2.69	0.0086	0.0852
Calendar ConstructionUpdateV2_11	-0.0347099	0.006148	-5.65	0.0000	0.2901
Calendar ConstructionUpdateV2_12	0.0300888	0.003589	8.38	0.0000	0.4740
LEU28 MOVEMENT QUANTITY_IN_100KG/1000_14	2.85427	0.7019	4.07	0.0001	0.1749

sigma	0.229386	RSS	4.10418568
R^2	0.602866	F(5,78) =	23.68 [0.000]**
Adj.R^2	0.577408	log-likelihood	7.59916
no. of observations	84	no. of parameters	6
mean(LRM01Ev2)	8.91257	se(LRM01Ev2)	0.352863

AR 1-2 test:	F(2,76)	=	0.48785 [0.6159]
ARCH 1-1 test:	F(1,82)	=	0.56582 [0.4541]
Normality test:	Chi^2(2)	=	2.7765 [0.2495]
Hetero test:	F(10,73)	=	0.78594 [0.6421]
Hetero-X test:	F(20,63)	=	0.81004 [0.6924]
RESET23 test:	F(2,76)	=	0.47144 [0.6259]

Source: Estimated by the author.

To test whether regression coefficients of the model are statistically significant, a t-test is applied. The t-test is for individual variables to test whether the coefficient of the variable differs statistically significantly from zero or not. Test process is done in the following way: a coefficient is divided by its standard error, which will result the t-value (equation 12). The t-value is then compared to Student's t-table with size of sample and with a selected confidence level, 5 % is the usual comparison but also the 1 % level is monitored. The t-test is for a single regression coefficient, whereas an F-test is a test for all coefficients at the same time. Formula to calculate a t-value is the following (equation 12):

$$t - value = \frac{\beta_1}{\sigma_1}, \tag{12}$$

where β_1 is the regression coefficient of the variable estimated and σ_1 is the corresponding standard error for the specific variable. Hypothesis for the t-test are:

H0: Coefficient = 0

H1: Coefficient ≠ 0.

(Verbeek 2008, 25.) The critical value for this sample size at 5 % confidence to reject H0 hypothesis and to accept alternative hypothesis is 1.99. Corresponding critical value at 1 % confidence level is 2.66. From the test results reported in the table 5 one will find out that all t-values of regression coefficients are above critical value of 1 % confidence level when the test result is positive or are below the critical value of 1 % confidence level when the test result is negative. Other way to interpret test results is to observe t-prob value, which indicates the probability similarly as the confidence level points it. When one has 6% t-prob value, it will imply that there is six percent probability to make a false conclusion, i.e., to reject H0 hypothesis and accept H1 hypothesis. In the case of this model, one can conclude that there is a marginal change to make a false conclusion. One should reject H0 hypothesis for all explanatory variables and accept the alternative hypothesis that the coefficients are not zero.

After one has estimated the regression model and has the test results, to have statistically sound model one must test the error term, the residual. In residual testing, one tests whether distributional assumption can be applied or not. With OLS estimation, one will have different residual tests to make sure that the residual is well behaving. Residual, the error term, is calculated from the deviations between the model and the actual time series. OLS estimation will automatically fit the best possible model that minimizes the sum of least squares, hence the name of OLS. Due to this methodology, one should then test whether that error term has all the required properties for a sound model. Residual test reported in the table 5 are autoregressive (AR) test, test for autoregressive conditional heteroscedasticity (ARCH), normality test to test whether residual follows a normal distribution, heteroscedasticity tests and model specification, the Ramsey RESET23 test.

In order to trust the estimated OLS regression model, Gauss-Markov conditions must be met. Gauss-Markov conditions that justify the usage of OLS estimator are the following:

$$E\{\varepsilon_i\} = 0, \quad i = 1, \dots, N \qquad (A1) \tag{13}$$

$$\{\varepsilon_i, \dots, \varepsilon_N\}\ and\ \{x_i, \dots, x_N\}\ are\ independet\ (A2) \tag{14}$$

$$V\{\varepsilon_i\} = \sigma^2, \quad i = 1, \dots, N \qquad (A3) \tag{15}$$

$$cov\{\varepsilon_i, \varepsilon_j\} = 0, \quad i, j = 1, \dots, N, \quad i \neq j \qquad (A4). \tag{16}$$

Gauss-Markov assumptions should be read that the expected value $E\{\}$ of error term ε_i, the $E\{\varepsilon_i\}$ must be zero (A1) (equation 13). In the long-run, i.e., on average the linear regression line that the model presents should be correct. The second assumption (A2) (equation 14) states that the residuals ε_i and actual values x_i are independent. The third assumption (A3) (equation 15) states that the residual is homoscedastic, not heteroskedastic. The last assumption (A4) (equation 16) is that there is no correlation between the residuals. This correlation is later stated as autocorrelation in the error term. (Verbeek 2008, 16.)

Autoregressive test is applied in order to test, whether there is autocorrelation in the residual. Alternative naming for this is error autocorrelation test, to highlight the fact that one is testing the error term. Both positive and negative autocorrelations are in the scope of test and interest. If there is positive autocorrelation, the negative deviation is followed by a negative deviation and vice versa.

With negative autocorrelation, the next deviation will have the opposite value; positive deviation will be followed by a negative deviation. Positive and negative autocorrelation are illustrated in the figure 6.

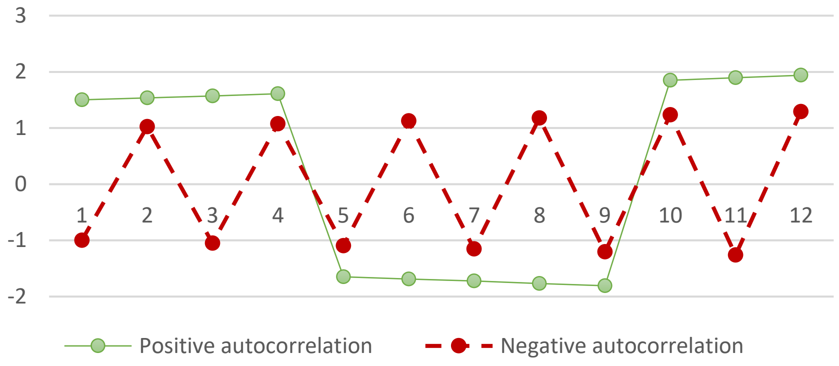


Figure 6. Understanding autocorrelation (estimated and complied by the author).

With autocorrelation test, one is testing whether Gauss-Markov assumption (A4) is valid (Verbeek 2008, 16). Formula (equation 17) for the autoregressive (AR) test is the following:

$$u_t = \sum_{i=p}^{r} \alpha_i u_{t-i} + \epsilon_t, \tag{17}$$

where $0 \leq p \leq 1$. In large sample values are $\chi^2(r)$ distributed. (Greene 2012, 949-950.) For small sample testing, the F-form of the test is used and which is applied as well in this thesis (Harvey 1990). Hypothesis for the autoregressive (AR) test are:

$$H0\text{: } \alpha_i = 0, \textit{ residual is not autocorrelated,}$$

$$H1\text{: } \alpha_i \neq 0, \textit{ residual is autocorrelated.}$$

Autocorrelated residual is unwanted feature for the model. If there is autocorrelation of any sort, then the deviation from zero can be forecasted by using the previous deviation. When residual is autocorrelated one cannot trust the forecast since the deviations are not random, i.e., are not white noise anymore. To assume normal distribution, the residual must be white noise. (Greene 2012, 950.) The test result of the autoregressive test reported in the table 5 for the residual is F(2,76) = 0.48785 (0.6159). The value in brackets indicates the p-value of the test measure in percent. The critical value for F(2,76) in 5 % significance level is between 3.07 F(2,120) and 3.15 F(2,60). The test value is not even close to the critical value, test fails to reject the null hypothesis and will one concluded that there is no autocorrelation present in the residual and that the model is statistically sound on the behalf of autocorrelation.

When a time series has a time depending variance, so that the historical variance determinates the future variance of the time series and when one can express that dependence by an autoregressive model, then one has an Autoregressive conditional Heteroscedasticity (ARCH) effect in the time series or in the model. When the ARCH prevails, one can use the historical stationary observations to estimate the future conditional variance. (Heij 2004, 621-622.) In this thesis, the ARCH effect is not desired feature since OLS do not accept the ARCH to be present because the residual will not be well behaving and the forecast cannot be done with usage of the normal distribution. When the ARCH effect is valid, it should be modeled with ARCH or GARCH models. Formula (equation 18) for the ARCH test is the following:

$$E[u_t^2 | u_{t-1}, \dots, u_{t-r}] = c_0 + \sum_{i=1}^{r} \gamma_i u_{t-i}^2, \qquad (18)$$

where c_0 is the constant and then in the sum function one has residuals u_i^2 from the original regression model. Residuals are in the power of two since the OLS estimator will minimize the sum of least squares. The test is whether the coefficient of the residuals differs statistically from zero or not. The test formula can be presented also by using the R^2, which is then multiplied by the size of the sample, T, to get TR^2. The ARCH test is asymptotically $\chi^2(r)$ distributed. The test can be presented in F-test form as well, which is reported in this case. The hypothesis for the test are:

H0: $\gamma = 0$, residuals do not have ARCH effects,

H1: $\gamma \neq 0$, residuals do have ARCH components.

(Engle 1982.)

Test statistic for the F-test is F(1,82) = 0.56582 (0.4541) as reported in the table 5. The value in brackets indicates the p-value of the test measure in percent. The critical value for F(1,82) in 5 % significance level is between 3.92 F(1,120) and 4.00 F(1,60). The test value do not exceed the critical value and the test fails to reject the null hypothesis. Due to this, one must accept the null hypothesis, conclude that there is no ARCH effect in the residual and the model is statistically sound on behalf of the ARCH.

To test whether the residual is normally distributed, one could test that with a method introduced by Doornik and Hansen in (1994). For the (equation 19), let one denote μ, σ_x^2 as the mean and the variance of $\{x_t\}$. The mean is then written as $\mu_i = E[x_t - \mu]^i$ and the variance correspondingly as $\sigma_x^2 = \mu_2$. Skewness and kurtosis (equation 19) of the distribution are denoted as

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} \text{ and } \beta_2 = \frac{\mu_4}{\mu_2^2}. \tag{19}$$

For a sample, the corresponding (equation 20) counterparts are:

$$\sqrt{b_1} = \frac{m_3}{m_2^{3/2}} \text{ and } b_2 = \frac{m_4}{m_2^2}, \tag{20}$$

where the m_i replaces the formerly denoted μ_i. (Doornik & Hansen 1994.)

When distribution is normal, it has skewness of 0.03 and kurtosis of 2.96. The skewness and kurtosis are important to have, since those will indicate whether the distribution of the residual deviates from the normal distribution. Doornik and Hansen (1994) test statistic is derived from Shenton and Bowman (1977) who introduced for b_2 the usage of gamma distribution. Doornik and Hansen in (1994) applied D'Agostino's (1970) contribution that the distribution of $\sqrt{b_1}$ should be estimated by the Johnson S_u. When all of this is applied and $\sqrt{b_1}$, b_2 are transformed to z_1^2 and z_2^2, can one eventually test the normality assumption. The normality test static (equation 21) can be calculated in the following way:

$$e_1 = z_1^2 + z_2^2 \sim \chi^2(2). \tag{21}$$

Hypothesis for the normality test are:

H0: Residual is normally distributed,

H1: Residual is not normally distributed.

To use the normal distribution for forecasts and to have Gauss-Markov assumptions fulfilling model, the residuals must be normally distributed. When residual is normally distributed, one will have normal distribution at disposal to calculate possible values around the forecast. Test statistics for the χ^2 normality test is $\chi^2(2) = \mathrm{Chi\char`^2(2)} = 2.7765$ (0.2495) as reported in the table 5. The value in brackets indicates the p-value of the test measure in percent. The critical value for $\chi^2(2)$ is 5.99 in 5% significance level. The test value do not exceed the critical value in 5% significance level and the test fails to reject the null hypothesis. Due to this, one must accept the null hypothesis, conclude that residual is normally distributed and the model is statistically sound on the behalf of normal distribution assumption.

When one has varying variance in the error term ε_i at the same time when residuals are uncorrelated with each other, the residual is heteroscedastic. When the residual is heteroscedastic, the diagonal $V\{\varepsilon|X\}$ do not equal the variance σ^2 times the identity matrix. To test whether there is heteroscedasticity in the residual, one can use White test (1980) to test it. When performing test for heteroscedasticity, one is testing whether Gauss-Markov assumption (A3) is valid in the model or not (Verbeek 2008, 16). If the residual is homoscedastic the estimator (equation 22),

$$\hat{V}\{b\} = s^2(\sum_{i=1}^{N} x_i x_i')^{-1}, \tag{22}$$

will give consistent estimation of the actual $V\{b\}$. Highly useful aspect of White's heteroscedastic test is that it do not require a determination of the type of heteroscedasticity one might experience in the residual. To test appearance of heteroscedasticity, one should calculate (equation 23),

$$NR^2, \tag{23}$$

which is F-test equivalent, where N is the size of sample and R^2 comes from the model results. For heteroscedasticity the hypothesis are:

H0: Residual is homoscedastic

H1: Residual is heteroscedastic.

The Hetero-x is the same test but cross-products are added to the test formula. (Verbeek 2008, 99.) Test for heteroscedasticity is done in two parts. First, the residual is tested on the behalf of heteroscedasticity without cross-products and then cross-products are added.

Test statistics reported in the table 5 for the F-test of homoscedasticity is F(10,73) = 0.78594 (0.6421). The value in brackets indicates the p-value of the test measure in percent. The critical value for F(10,73) in 5 % significance level is between 1.91 F(10,120) and 1.99 F(10,60). The test value do not exceed the critical value and the test fails to reject the null hypothesis. Due to this, must one accept the null hypothesis and conclude that residual is homoscedastic. Test statistics for the F-test of heteroscedasticity with cross-products is F(20,63) = 0.81004 (0.6924). The value in brackets indicates the p-value of the test measure in percent. The critical value for F(20,63) in 5 % significance level is 1.75 F(20,60). The test value do not exceed the critical value and the test fails to reject the null hypothesis. Due to this, must one accept the null hypothesis and conclude that residual is homoscedastic.

Both heteroscedasticity tests fail to reject null hypothesis and one must conclude, that the residual is homoscedastic and the model is statistically sound on the behalf of homoscedasticity assumption.

In (1969) Ramsey suggested a test for model specification testing to replace or to complement other model specification tests. The upside of Ramsey (1969) RESET test is that it do not take an opinion about the alternative hypothesis, the alternative hypothesis do not need to be specified exactly. If the model is correctly specified then the higher power elements of the auxiliary regression should have zero coefficients. To test whether the higher power variables have coefficients deviating from zero, following regression (equation 24) is formulated:

$$y_i = x_i'\beta + \alpha_2\hat{y}_i^2 + \alpha_3\hat{y}_i^3 + \cdots + \alpha_Q\hat{y}_i^Q + v_i. \quad (24)$$

F-test is then applied to determinate, whether one fails to reject null hypothesis or must one accept the alternative hypothesis. Hypothesis for the RESET test are:

H0: $\alpha_2 = \cdots = \alpha_Q = 0$*, Model is correctly specified,*

H1: Model is misspecified.

(Verbeek 2008, 66.) The test result of the model specification RESERT23 test for the residual reported in the table 5 is F(2,76) = 0.47144 (0.6259). The value in brackets indicates the p-value of the test measure in percent. The critical value for F(2,76) in 5 % significance level is between 3.07 F(2,120) and 3.15 F(2,60). The test fails to reject the null hypothesis and one will concluded that the model is correctly specified and the model is statistically sound to be used.

With an F-test, one tests whether regression coefficients of the model are zero at the same time or is at least one coefficient deviating from zero while others are zero. Hypothesis for the test are:

H0: All regression coefficients are zero at the same time,

H1: At least one regression coefficient is not zero at the same time as others are.

Formula for F-test (equation 25) is the following and normally econometric packages report the test measure automatically.

$$F = \frac{(S_0 - S_1)/(K-1)}{S_1/(N-K)}, \quad (25)$$

where S_0 is the sum of squared residuals from the restricted model and S_1 is from the full model. S_1, sum of squares from the full model, is calculated as $S_1 = \sum_i e_i^2$ and since the restriction model is that all the coefficient of the model are at the same time equal to zero, one will have only the intercept term in the restriction model hence $S_0 = \sum_i \ (y_i - \bar{y})^2$. In (equation 25), K is the number of regressors and N is the size of the sample. The idea is to test, whether restrictions have an influence to the test results or not, do the sum of squares differ from each other. The test can be presented (equation 26) in another notation as:

$$F = \frac{(R^2)/(K-1)}{(1-R^2)/(N-K)}. \tag{26}$$

(Verbeek 2008, 28-29.) F-test is performed in this case with five variables and with sample size of 78. Sample size for the test is derive by reducing the count of variables and constant in the model from the actual sample size 84, i.e., 84 - 5 - 1 = 78. Critical value for F-test is from F-distribution; in this case, it would be 2.332 in 5 % significance level and 3.261 in 1 % significance level. Since the test result reported in the table 5 for F(5,78) test is 23.68, one can reject H0 hypothesis with high confidence and accept the alternative H1 hypothesis and conclude that not all regression coefficients are zero at the same time.

3.5 Interpret the model

In the model, there are variables with different scales, which will enforce reader to pay attention when interpreting the regression coefficients of variables. Since the dependable sales variable is as natural logarithm, change of it will always be in percent. Interpreting natural logarithm will imply that the variable will increase or decrease X amount of percent. Explanatory variables are in two different scales: the construction and manufacturing are as an index but will be interpreted as units. This will mean that change in variable is either increase or decrease in units but the unit is from an index where 2010 in the baseline, level 100. Explanatory variable movement is as well as natural logarithm and due to this: it will change percent as well similar to the sales. There are four different possible cases for coefficient interpretation and these are (equations 27-30):

$$1)\ y_t = \alpha + \beta_1 x_{t-1}, \tag{27}$$

$$2)\ \ln(y_t) = \alpha + \beta_1 x_{t-1}, \tag{28}$$

$$3)\ y_t = \alpha + \beta_1 \ln(x_{t-1}), \tag{29}$$

$$4)\ \ln(y_t) = \alpha + \beta_1 \ln(x_{t-1}). \tag{30}$$

In the first case (equation 27), when x_{t-1} increases one unit, the y_t will increase β_1 units. In the second case (equation 28), when x_{t-1} increases one unit, the $\ln(y_t)$ will increase $\beta_1 * 100$ percent. In the third case (equation 29), when $\ln(x_{t-1})$ increases one percent, the y_t will increase $\frac{\beta_1}{100}$ units. Then in fourth case (equation 30), when $\ln(x_{t-1})$ increases one percent, the $\ln(y_t)$ will increase β_1

percent. The second (equation 28) and fourth (equation 30) possible variable forms are in question with the forklift model. One thing to note, the change in the dependable variable is an expected change in that variable.

Now that the model is ready, the next step is to interpret the coefficients of the variables in the model. Although one can interpret coefficient separately, one should bear in mind that monitoring and analyzing the model as a whole is crucial. Using two lags per variable can cause difficulties to interpret the model or there can be some sort of errors in the coefficients, e.g., one was expecting positive coefficient when using that lag only as an explanatory variable but in this model, the coefficient is actually negative. The high correlation of variables can and will cause difficulties to interpret the regression coefficients and one can observe somewhat conflicting results. The model is created to be the most accurate forecast model and not only to indicate relationship between two variables. In OLS estimation the regression coefficients are interpret as ceteris paribus. The ceteris paribus assumption means that only one variable will change at the time while others will stay as constant. Next, the coefficients are carefully interpreted.

When variable construction, lagged for 11 months, increases one unit, the expected change in sales is a 3.47 % decrease. In the case of the 12 months lagged construction variable, when construction increases one unit, the expected change in sales is a 3.00 % increase. This is interesting that based on the model: the same change will have different effect to sales depending of time. One could interpret this contradiction of increase and decrease, that 11 months earlier construction firms have already made investment decisions and are now focusing on their core business. 12 months ago situation in construction business was booming and encouraged the investment decisions of firms and that is visible through the regression coefficient as an increase in both. However, then 11 months from now, the change in explanatory will imply decrease in dependable sales variable. This might be, because the construction business starts to increase and the machinery ordered earlier are already delivered in the previous period, which will seem as a decrease. Since forklift trucks are not sold constantly and as a steady flow, there is cyclicality in sales and naturally, the customers will require the deliveries of the machines to suit their timetables. If there is not enough work to perform with the new forklift truck, why should one take the delivery then, why not wait for the seasonal boom to start and then use the machine.

Another important factor for cyclical year-end sales, not only internal sales behavior to show investors high results, but also the customers will need the machine to be able to deduct the depreciation in taxation. It is more rational to take the delivery of machinery in the later part of the year than yearly next year; one can deduct full deprecation in taxation immediately and lower the profit under income taxation, i.e., pay less tax for the same profit. This tax benefit is applied in some EMEA countries, e.g., in Finland, Sweden and in UK but not, e.g., in Germany or in France as discovered earlier.

When variable manufacturing, lagged for 11 months, increases one unit, the expected change in sales is a 3.00 % increase. In the case of the 15 months lagged manufacturing variable, when the manufacturing increases one unit the expected change in sales is a 1.30 % decrease. There is a difference between the construction and the manufacturing when both are in the same model. The 11 months lagged manufacturing and sales will change to same direction, an increase will cause an increase. However, with the construction, the direction was the opposite for the 11 months lagged value: an increase is followed by a decrease and vice versa. The 15 months lagged manufacturing causes a change to opposite direction as it changes, increase will cause decrease and vice versa.

When all these variables are in the model, one could interpret that either the investment or the reaction time for investment needs or the investment process completion is shorter in manufacturing than is in the construction industry. This could cause that increase in manufacturing 11 months ago increases the sales 11 later, which could be the timeframe for equipment deliveries. 15 months lagged variable has an opposite change to sales, increase in variable will cause decrease in sales. This could be explained by that the customers using the forklift trucks focus on their core business, manufacturing, and their production increases but the equipment for that increase are already delivered. It would be reasonable, that four quarters earlier, firms start their investment process based on the environment at that moment to prepare one year ahead. One quarter earlier same firms where full-on production phase and there were not enough equipment to deliver, i.e., the sales will decrease.

The fifth variable in the model is the movement of goods in kilos, which is lagged for 14 months. Movement has a natural logarithm transformation for combined imports and exports in hundreds of tons in kilos. This natural logarithm transformation will imply that when goods movement increases one percent the sales are expected to increase 2.85 percent. This if intuitive since goods are lifted and transported at some point or in multiple points of supply chains by forklift trucks and the moved quantities are large. One will have millions of kilos of goods transported each month from EU to outside EU and vice versa. One percent increase in kilos should large increase in sales as well but not a massive one since Kalmar is not the only forklift truck manufacturer in the EU28 area.

When one wants to test the fit of the model to the actual data, one will derive R^2 (equation 31) and variable count adjusted $\bar{R}^2$. The normal interpretation of R^2 is how much in terms of percentage the model is capable to explain of variation of the dependable variable. Normal boundaries for R^2 and later $\bar{R}^2$ are 0% and 100%, i.e., if the model can explain perfectly the dependable variable, it will have a value 1 = 100%. In some special cases, one can have a negative R^2 value.

$$R^2 = corr^2\{y_i, \hat{y}_i\} = \frac{(\sum_{i=1}^{N}(y_i - \bar{y})(\hat{y}_i - \bar{y}))^2}{(\sum_{i=1}^{N}(y_i - \bar{y})^2)(\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2)}, \tag{31}$$

where one has squared coefficient correlation from the sample between the actual values and the fitted values derived by the model. The drawback when using R^2 is that it will never decrease if one adds more variables. To "punish" of this and to have small amount of regressors in the model, adjusted R^2 (equation 32) is introduced.

$$\bar{R}^2 = 1 - \frac{1/(N-K)\sum_{i=1}^{N} e_i^2}{1/(N-1)\sum_{i=1}^{N}(y_i - \bar{y})^2}, \tag{32}$$

where N is the sample size and K is the amount of regressors in the model, to punish of an excess usage of variables. When one uses the R^2 or the $\bar{R}^2$ together with OLS estimation, these measures will indicate for the researcher how well the model fits to the data. Especially with the adjusted $\bar{R}^2$, the punishment effect will create a situation where it might not increase even though one adds one more variable. Using the adjusted $\bar{R}^2$ forces the researcher to look for the most efficient model and not to include extra variables to the model. (Verbeek 2008, 23.)

3.6 Conclusions of the model

From the statistical analysis perspective, the model has all elements to be concluded to be a statistically sound and one can trust the results it presents. All variables in the model are $I(0)$ and therefore any linear combination will be $I(0)$. The model can be used to forecast future values of the dependable variable by using lagged values of the explanatory variables only. The residual of the model is well behaving, all statistical test are passed about the behavior of the residual, there is no autocorrelation in the residuals, the residual is ~ normally distributed, residual is homoscedastic and do not have ARCH effect, and the model is correctly specified. One thing to note is that with these descriptive tests all test are passed with excellence. One might add further variables to increase R^2 even though it might decrease some of descriptive test statistics.

Since the model is constructed using the lagged values of explanatory variables, one can forecast future with using only lagged values of explanatory variables. The latest lagged value is almost from one year earlier, which is highly useful for financial forecasting in business life. The model can truly complement the normal budgeting process and a company to plan its operations as efficiently as possible and to focus on key areas and possible weak spots. Variables used in the model are fair and sound by common sense and statistical analysis prove that as well. Construction and manufacturing are important fields where one operates with forklift trucks. Goods movement fulfills the model since the actual goods are as important as are the industries where forklift trucks are used.

A caveat in model is the usage of dual lag per variable, which can produce contradicting regression coefficients, but one should bear in mind that the model as a whole is the one that matters. Other detail is the usage of eleven lags. For example, if after normal summer holiday season, somewhere in August, one starts budgeting process for the next year will that yield a need for over one year explanatory variables' values. That would imply that in order to forecast the whole next year, one would need data for 16 months. Naturally, one will have the maximum of 11 months of data to use from history, i.e., one will need forecasts for explanatory variables as well to have enough values for explanatory variables to forecast the next calendar year. Forecasting from expected values can be done but it will not be the most accurate way to proceed. Even though there are some aspects that do not make the model perfect, the upside is that the model suits perfectly for rolling forecasting.

Even if the model is not a compatible with static budgeting where one performs a budget for next year, with this model one will make forecast along the way but fact based. Seeing eleven months ahead could give a company time to react and should make it more agile. Being more agile and future oriented should in the long-run also benefit the shareholders. The company would be able to deliver results when it is possible and improve weaknesses of supply chains, invoicing and other internal issues. With forecast for eleven periods, a company might even have enough time to adjust the production or the capacity in order to increase profitability. If the forecast indicates lower sales for a specific month, this could be taken into consideration well in advance and other activities could be invented either to increase sales or to reduce costs in that month.

To have more flexibility for the production and for the fixed costs, company might use flexible working hours where workers will work extra when the production is at full stint and would the use the working hour bank to have time off when there is no need for all workers to be at the production facility. With this arrangement, the company can decrease the labor costs, since the salary for workers is the same but the working hours flex based on the production needs. This arrangement was implemented at Nokian Tyres Plc. (Nokian Tyres 2018.) If one can have such arrangements for production facilities or in other flexibilities, it will benefit the company in the long-run.

The R^2 value of the model is high enough that the forecasted values will not have exceedingly high standard errors. A high standard error implies wide distribution of possible values when one calculates those from the normal distribution. Naturally, the high R^2 should increase the accuracy of the forecast value alongside the distribution of the forecast. One thing the high R^2 will not help is forecast accuracy when there is a large amount of shocks that are not explained by fundaments but which are entirely finance related: whether company were able to deliver and invoice machines as planned, whether customers declined of the delivery in certain month and postponed to another month. All these can affect the accuracy and some of those effects might be still outside of the model. One might lack a variable or variables from the model that describe these effects accurately.

For the future development, one could model the customer behavior, which might result a good insights to invoicing and cyclic sales behavior. The odd behavior of sales must be explained with entirely different variables than the market fundaments that are in the scope of this thesis. Beside customer behavior, one could add and model internal behavior whether there are system errors or other reasons why the delivery did not happen even though the customer would want to receive the machine.

4 Financial forecasting

4.1 Forecast for year 2018

When one performs a forecast using the model, one will have a forecast for next 11 months. At the time of writing this thesis, the forecast can be done from end of 2017 for 11 months forward, i.e., until the end of November 2018. Forecast for the next 11 months is presented in the figure 7. The forecast is estimated using a parameter uncertainty. With parameter uncertainty, one will have different standard errors for each forecasted value. This will yield different results than using error variance only. The usage of parameter uncertainty is clear since some forecasts can be quite accurate but with some forecasts there can be issues and the distribution will be marginally wider than with the other more accurate forecasts. Uncertainty can rise if the model cannot capture the behavior of certain internal or customer related issues that are not cause by the industry fundaments.

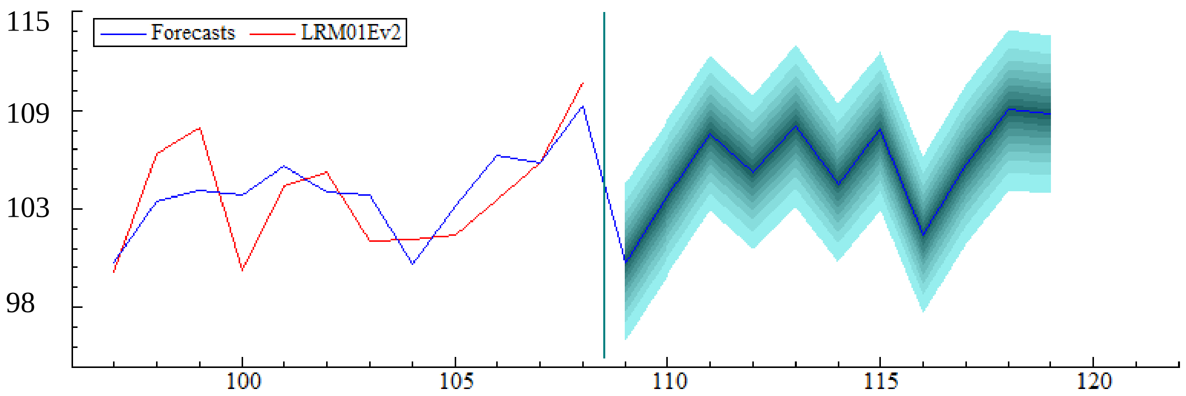


Figure 7. Forecast for 01 – 11 / 2018 periods (estimated by the author).

From the figure 7, one will notice that the sales forecast for the next 11 months is not stable. The model captures the true nature of the sales behavior and interpolates it forward. January has a low forecast value similar to August. For January, the reason could be that there is no need for deliveries for large amount of forklift trucks. August could have a similar reason behind it; August is the time for holidays in Central Europe. When children are not in school, parents will have their vacation as well. When people are on vacation, there is not that much activity in the industry and no need to deliver new equipment since nobody will be there to receive those or to use the new equipment. After the holiday season, the activity in industry wakes up and the forecast has increasing trend towards the end of the year.

Forecasted values for next 11 months are reported in the table 6. For the forecast estimates, there are the corresponding standard errors for the estimates.

Table 6. Dynamic ex ante forecasts for LRM01EV2, natural logarithm transformation of the sales, and standard error for forecasts with parameter uncertainty. Indexed to January 2018, base level 100.

Horizon	Forecast (indexed Jan)	Standard Error
January	100.00	0.2416
February	103.76	0.2397
March	107.49	0.2394
April	105.24	0.2371
May	107.90	0.2502
June	104.53	0.2422
July	107.69	0.2443
August	101.59	0.2413
September	105.66	0.2434
October	108.84	0.2463
November	108.57	0.2411

Source: Estimated by the author.

One will see that there is more uncertainty associated with the forecast for January than there are with forecasts for February, March and April. What is strange with the standard errors is the standard error for May. May is not the end of quarter nor the end of half-year month, and it has the highest standard error of all the 11 forecasted months. From June onward the standard error is stable and lower than in May. The last forecast, for November, has fourth best standard error even though normally the uncertainty of forecast will increase the further from starting point one will go. The reason for this could be the large uncertainty of revenue recognition during the middle of the year. At the year-end, everything possible will be invoiced and the forecast could be because of that more accurate than in the summer periods.

Based on actual sales development for 2018, the forecast model do perform once again with high accuracy. For January, the estimate was too low compared to actuals but the estimate for sales in February was spot on. The operations forecast and the budget were much higher than the actuals or the forecast model’s estimate for the first two months.

One thing to mention is the usage of logarithm variable as a dependable variable. Since the logarithm scale is not linear, the higher logarithm values will transform to higher amounts in Euros. With this is meant that when one transforms logarithm values to numeric, there will be more probability mass above the forecast than below it for same probability since the scale is not linear. Moreover, due to this, the end of the year values will have much wider spread for the monetary values for the same 90 % probability distribution. In the figure 8 is presented the forecast with 90 % probability distribution around indexed forecast. Indexing was done using the natural logarithm values and due to this, the figure 8 is not affected by the transformation.

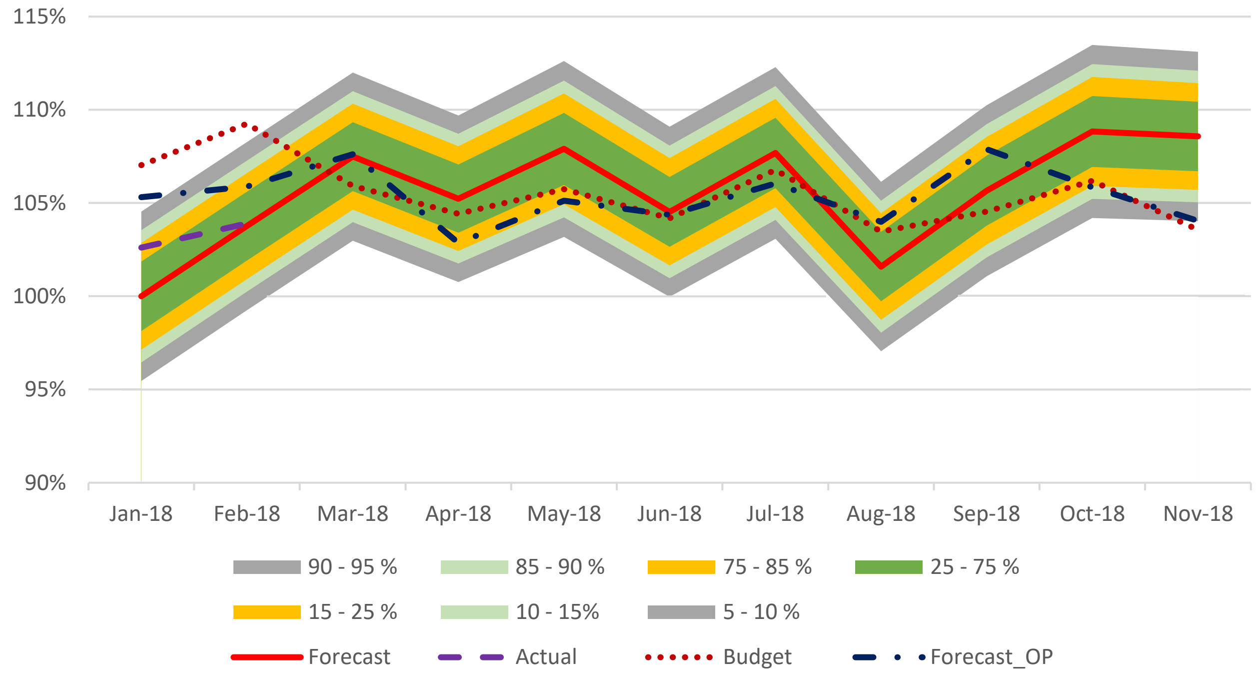


Figure 8. 90 % profitability distribution for the 11 forecasted values. Forecast for 01/2018 is the index level 100 (estimated by the author).

When observing and analyzing the historical data, there is a clear change in the forecast pattern for 2018. The forecast model indicates more sales for spring and summer months, from March to July. It seems that there is a clear step up in sales. The upside of this indication that it is completely derived from the market fundaments and from that perspective, one could and should trust it.

Comparison of forecast is now performed against the budget and the most up-to-date forecast, both of which are made by the same persons. Budget is the financial plan for next year and the budgeting process normally starts after the summer holiday season. Even though the budget is somewhat old, it is the most precisely made forecast. The most up-to-date comparison series is the Forecast_OP, which has been made in January to have the most comprehensive understanding of the sales development for the rest of the year.

From the figure 8, one will see that for January the forecast model did not capture accurately the sales but did for February. For January, there was roughly 20 % probability for that particular sales level to happen but more interestingly, based on the forecast model there was only a marginal change that the sales volume indicated the budget or by the Forecast_OP to happened. For the mid-part of the year the budget and forecast model do seem to follow each other but interestingly, the updated Forecast_OP does not. In addition, especially the end of year hockey stick (Chen 2000, 186) effect is not budgeted at all.

4.2 Forecast for year 2017

To test how the model actually performs, one can estimate the model without all values of dependable variable. For example, in this analysis the model is estimated until the end of 2016. After that, one can then "forecast" next 12 months but using the regression coefficients that would have been estimated at the end of 2016. With this is shortcut one will have the correct coefficients, which one would have had back in 2016. From statistical perspective, there is no bias for the test since the future actual values will not affect the regression coefficients. The forecast is easy to process since all the explanatory variables have data. One extra aspect of testing the model in this way is to benchmark against operations personnel' budget forecast. As previously, the forecast is done with parameter uncertainty to have varying standard error for forecasts and not only one static error.

One issue comparing to budget might be incentive-based manipulation for the budget. Walker and McClelland (1991) found that sales operatives might boost the sales budget in order to balance between two different and conflicting objectives. First, the sales organization must have high sales to justify their allocated costs; the profitability of the part of organization must be high enough. Secondly they must motivate and encourage personnel via incentive plans to perform better, the targets must be able to achieve, i.e., the target sales cannot be too high. (Walker & McClelland 1991, 379.)

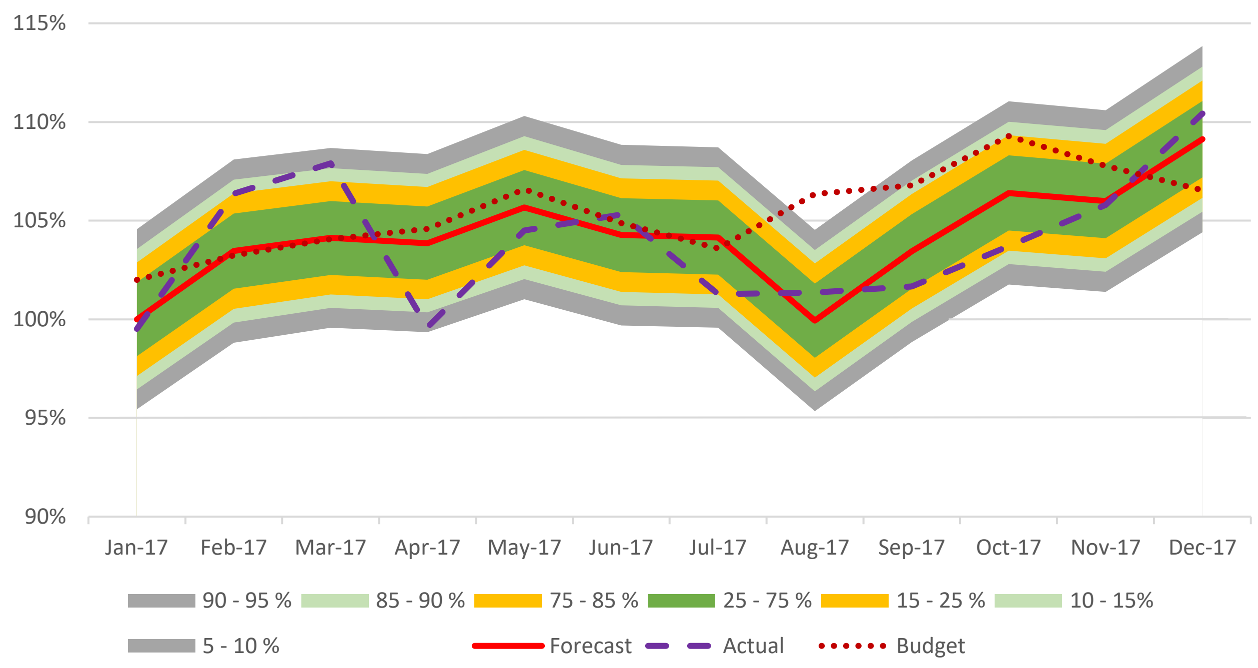


Figure 9. 90 % profitability distribution for 2017 forecasts. Forecast for 01/2017 is the index level 100 (estimated by the author).

From the figure 9, one will immediately notice the same August effect that was visible in the forecast for 2018. In addition, one will notice the increasing year-end sales effect, so called hockey stick – effect (Chen 2000, 186). From figure 9, the difficult start for the year is clearly visible; probably customers will not need new forklift trucks in the middle of winter. Naturally, there are areas in EMEA that do not suffer from winter but there are many highly snow sensitive areas, e.g., Sweden, Finland and even Germany occasionally. These snowy areas are a small proportion based on count but do present important part of the total sales and the winter is due to that a plausible factor to explain the dropping sales in addition to the hockey stick effect (Chen 2000, 186) in December.

Table 7. Mean and standard error for the forecasted periods

Horizon	Forecast	SE	Actual	Error	t-value	-2SE	+2SE
January	104.32	100.00	100.00	-4.32	-0.175	64.31	169.20
February	77.64	101.82	100.00	22.36	1.028	47.44	127.03
March	71.9	99.88	100.00	28.10	1.366	44.36	116.54
April	145.22	98.97	100.00	-45.22	-1.559	89.99	234.35
May	110.6	101.99	100.00	-10.60	-0.409	67.54	181.13
June	91.35	100.33	100.00	8.65	0.373	56.22	148.41
July	128.36	100.29	100.00	-28.36	-1.030	79.03	208.5

August	88.3	100.79	100.00	11.70	0.511	54.23	143.77
September	116.81	100.99	100.00	-16.81	-0.636	71.67	190.38
October	126.6	101.7	100.00	-26.60	-0.959	77.43	207.00
November	101.79	101.03	100.00	-1.79	-0.073	62.44	165.93
December	89.29	103.52	100.00	10.71	0.453	54.13	147.24

Source: Estimated by the author.

From the table 7 is possible to see that two forecast do hit the target accurately. The forecast for November has t-value of -0.073, which equals a t-probability of 47%. Complete opposite for November is April, where the actual barely stays with-in the 90% probability mass. T-value for the forecast for April is -1.559 and with two-tailed t-distribution, the lower 5 % t-value is -1.645, the corresponding t-probability for the forecast is 6.11 %.

To determinate, what value the external fundamentals based forecast model will bring to everyday business life, it is important to benchmark against the operations personnel' budget. From table 8, one will see that the first half of the year, the forecast and budget are aligned and both are as much right as are wrong. The largest difference appears after July when the holiday season starts in the Central Europe and in some cases, people return to work in the Northern Europe. The budget is a lot larger than the forecast model implies the sales to be and during the autumn and early winter, the forecast model is performing much better than the budget does. One must remember that the budget is one year old, when the August actually comes but still, the performance of the budget is not adequate. The forecast model has better performance in the challenging part of the year and in fact in overall terms. Forecast model has a twelve-month forecast accuracy of 99.67 % and the budget has 89.86 % respectively. The deviation between forecast and the actual is roughly one unit of medium forklift trucks or two light forklift trucks.

Table 8. Comparison between forecasted, actual and budgeted sales

Horizon	Forecast (indexed)	Actual (index)	Budget (indexed)
January	104.32	100.00	124.19
February	77.64	100.00	76.13
March	71.90	100.00	71.51
April	145.22	100.00	154.62
May	110.60	100.00	119.97

June	91.35	100.00	96.30
July	128.36	100.00	122.31
August	88.30	100.00	154.52
September	116.81	100.00	156.34
October	126.60	100.00	162.76
November	101.79	100.00	118.81
December	89.29	100.00	71.27

Source: Estimated by the author.

To evaluate the performance of the forecast model, following key statistics are introduced. These are mean of the error term, standard deviation of the error term, mean square error (MSE), root mean square error (RMSE) and mean absolute percentage error (MAPE). (Mendenhall & Reinmuth 1993, 668-669.) Mean of the error indicates whether the residual has a mean of zero or not. One normal assumption in time economics is assuming normal distribution with zero mean and standard deviation of one. The residual of the model has a bit zero deviating mean, -0.02. This do not require further actions but is informative that the mean of the probability distribution is not exactly at zero. The standard error of the model is 0.20. To evaluate the performance of the forecast model, one could calculate following statistics:

$$MAD = \frac{1}{n}\sum_{t=1}^{n} |y_t - \hat{y}_t|, \tag{33}$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n} (y_t - \hat{y}_t)^2, \tag{34}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} (y_t - \hat{y}_t)^2,} \tag{35}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{y_t - \hat{y}_t}{y_t}\right| (100\%), \tag{36}$$

where y_t is the actual value from the series and $\hat{y}_t$ is the corresponding estimated value for the period. Mean absolute error (MAD) (equation 33) has a caveat for extreme deviations and due to this; mean square error (MSE) (equation 34) is introduced. MSE (equation 34) and the square rooted version RMSE (equation 35) will punish of the extreme deviations between the actual and forecasted value. The usefulness of this punishment or penalizing comes from the cost of wrong estimated value. If the result of incorrect estimate will endanger the organization then the model should be selected which has the best forecast capability. With mean absolute error (MAPE) (equation 36), one will have an issue of inflating the statistic. Since there is a quotient in the formula and the statistic is shown as percent, extremely low denominator values will increase the value of the MAPE (equation 36) extensively. (Mendenhall & Reinmuth 1993, 668-669.) Test statistics for the 2017 forecast are RMSE = 0.20473 and MAPE = 1.9217. Since the test values are to determine which model to choose, the test values itself do not imply that much information. In order to benchmark the values against a counterpart, corresponding test statistics are presented for 2016 forecast.

To test one-year forecast, the model is estimated up-to 2015 and then forecasted one year forward. Test statistics are RMSE = 0.23257 and MAPE = 2.0726. With same process is done for a two-year forecast and following test statistics are observed RMSE = 0.21768 and MAPE = 1.9738. With these test statistics, one will quickly see that the performance of the 2016 forecast was below the forecast accuracy for 2017. The forecast accuracy for two years is between 2016 and 2017 forecasts, which will imply the great performance of the model to forecast the 2017 sales. For further reading of the key statistic, one will find about MSE, e.g., Nahmias (2009), Hanke and Wichern (2009), Silver, Pyke and Peterson. (1998) and about the MAPE, e.g., Hanke and Wichern (2009).

4.3 Update frequency of the forecast model

Morlidge and Player (2010) discussed in their book about the update frequency for the forecast model. From their perspective, the frequency depends on two dimensions: about the variability and criticality. The variability refers how much there is variability to which the model should be calibrated and the criticality to the financial effect of that variability. Answer to the question is not easy to give but the update frequency should relate also to the possibility to make changes. When organization is committed to a financial target, there might not be enough flexibility in the organization to fulfill new commitments. In the beginning of planning for resources, the forecast model should be updated on monthly basis but when organization is in the execution mode, maybe every quarter is enough. They also identified and guided through the dilemma of too short forecasting capability. Their guidance is to use rolling forecasting and only forecast the amount of the longest lead-time of a company. With this is meant that if there is a 12-month lead-time that should also be the forecast horizon at all times. The normal accounting and calendar period forecasting is not fruitful forecasting period, since if there is the 12-month lead-time, the forecast horizon for the budgeting is longer, maybe 15 months. (Morlidge & Player 2010, 62-72.)

In order to test whether forecast model is usable with easy maintenance, the following forecast analysis is presented. Forecast is created by estimating the model until 2015. The idea of this test process is to test the coefficients of the model estimated and how those will perform in the long-run. Longer the horizon of which one can estimate the model more accurate are the coefficients. When the model performs well without estimating the model in every month, it will be highly useful in a corporate environment. Constantly updating forecast models is wasting recourses if the model performs accurately without constant recalibration. In a scenario, where models are estimated for example once in a quarter will that decrease the need for manual work and releases further possibilities to improve the models and not only to maintain already existing ones.

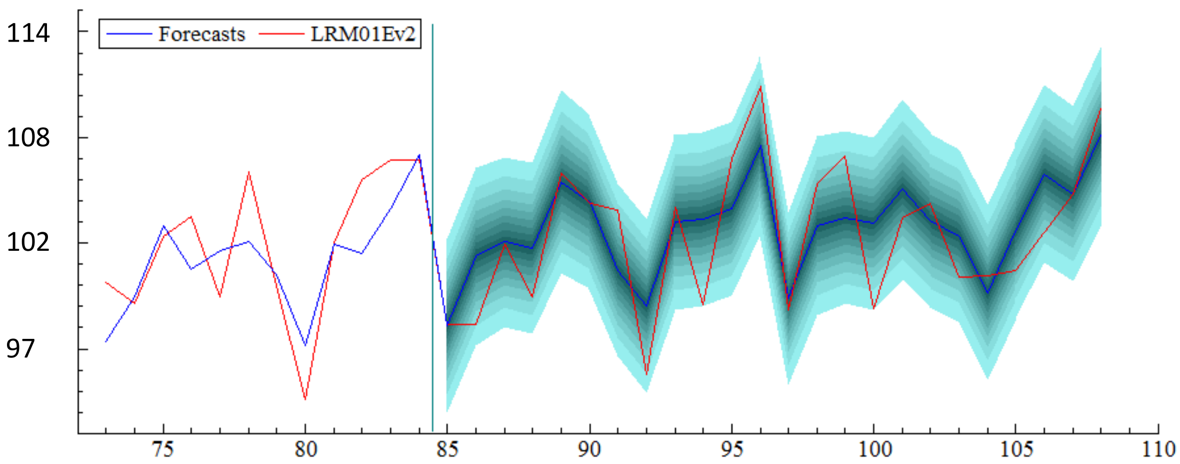


Figure 10. Forecast plot for a two-year forecast (estimated by the author).

From figure 10, will one notice that the forecast model do capture well the true development of the sales. All actual values are with-in the 90 % probability mass. At time point 100 (April 2017), the value is hideously close to exit the 90 % probability mass. April is one of those months, which the model in current setup do not work accurately. April as one will know by now is the month after Q1 closing. During every quarter-end, companies will have high pressure to deliver good results to the markets and maximize the good news about the company. As already mentioned, the Q4 is critical for delivering the results that the market is demanding but what if a company do not have good news to give in Q1. Most likely, in those cases they will try to gather all sales possible together during Q1. All of this will decrease again the future sales. The decrease in January is the natural result of calendar year based accounting period.

Thoughts by Morlidge and Player (2010) about the update frequency of the forecast model are valid. When one analyzes the visual presentation in figure 10, one will see the great performance of the model even thought it was not calibrated for two years. Furthermore, the fact that the forecast model has only a little historical data for the estimation, the performance is remarkable. With this in mind, there might not be need for constant calibrations but only when the crucial plans are about to beginning. One can see from figure 10 the same, which is presented in the table A in Appendix, when there is a drop that the model did not properly capture, the t-value will increase since the probability to encounter such value was rare. One should bear in mind that any internal issues or manipulations for the financial actions, which will affect the invoicing, are out of the scope of this thesis. These are furthermore analyzed later in the chapter 5.

4.4 Performance of the forecast model

Simply observable and at least to some extent important measurement of forecast performance is the count of correct forecasts. Forecast is correct when the forecast value is essentially correct and close to actual. Correct forecast are two out of 12 possible in the case of forecast for 2017 and eight out of 24 in the case of forecast for two years.

To predict exactly monthly revenue is difficult and in some cases impossible but to forecast the total revenue for the year is the one that matters the most. On yearly level, one will have the total costs, which will be experienced no matter when the machine will be delivered and invoiced, the manufacturing and invoicing are separate functions and are not linked to each other in that way. On yearly level the timing of sales revenue between months is not crucial as long as those are not postponed to another accounting period, i.e., to the next year. Incentives for invoicing and revenue recognition are problematic and to tackle those in the future, the model could be adjusted to capture the month specific invoicing behavior and customer behavior as well. For 2017, the forecast model outperforms the budget. The budget can contain target setting for sales, which the management has pushed there to have financial figures that board of directors are demanding. This bias has already discovered to happen in other companies and can be possible in this context as well. The most difficult question is to know whether there is this fixing done or not (Makridakis et al., 1998). Forecast for 2017 is extremely close to the actual value, the difference is one or two units, which can be then sold in 2018 or not be sold at all. By chance, one can be lucky once and because of that, the continuing performance is the one that matters. For 2016 forecast model indicates 100.41 (indexed) sales and the actual sales value is 100 (base index), again the forecast indicates one or two units more sales than there actually is, but is extremely close to the actual. Corresponding budget figure for 2016 was 93.49

(indexed). From this comparison, one will see the superiority of the forecast model versus static budget.

4.5 Conclusions of the forecast model and forecasts

To conclude the model and the forecasts derived from the model. The basis of the model are statistically and economically sound. The variables in the model are stationary and do not have autocorrelation. When variables are combined and model is estimated, the residual of the model is well behaving and one can trust the forecast that are derived from the model. The variable count adjusted $\bar{R}^2$ is high at 58%, which implies good accuracy for forecasting.

The model captures well the behavior and development of the sales time series. There are caveats in the model and there are room for improvements in the future. Important thing to note about the improvements is the fact the improvements will only improve the accuracy; the fundaments will be the same. This means that the improvement is needed because of internal issues to invoice machines or customers' reluctance for deliveries. The fundamental fact remains the same; the forecast indicated truthfully the actual sales value that should have been without the cyclic behavior and one should create a behavioral variable to include the non-fundamental effects. Creating such variable will be highly challenging but also rewarding is such is possible to create.

When one performs the forecast, the standard errors are small enough that the probability distributions of possible values do not increase extensively. The forecast model performs tremendously and the usage of it will bring value to Kalmar for a rolling budgeting process. Whether management is willing to rely on a statistical model is another question but the statistic foundation is solid and at least usage of it should open a dialogue about how to forecast. Model has excellent accuracy in terms of yearly sales forecasting and capturing the development pattern of sales. On yearly level, it is easier to match the sales and forecast since the deviation between actual and forecast are not always fundamental issues as already learnt but internal incentive derived issues or customer related delivery issues. In Kalmar, there will be the same challenges as discovered by other researches.

5 Extension to the research and conclusions

One question that rose frequently was the need to capture the cyclic non-fundament based behavior of the sales. After estimating the model and analyzing the fit of the model compared to the actuals, one will see the caveats. In the first sight, these seem to follow a certain pattern, e.g., after each quarter closing, half-year closing and year-end closing. Beside these normal stock market benchmarks, there are holiday season effects, e.g., in Central Europe in August. To address this issue, one could use dummy variables.

Dummy variable will be either one or zero, e.g., August would get one and the other months of the year would have a value of null. This could be applied to all end of quarter months, i.e., March, June, September and December would have value of one and otherwise zero. One could also create a dummy variable for the end of year effect, where December would have value one and other months would be zero. In theory, this sounds fair and straightforward process but the reality is different; the non-fundament sales behavior depends also from other factors. Even though the pattern of the non-fundament based sales behavior seems to be similar, there is much to understand the magnitude of the drops and peaks.

With the previous mentioned is meant that the drop is not always with same magnitude nor is the peak, all of these vary over the time. The challenge with dummy variables comes from the fact that each value of one assumes to have the same amount of effect. The dummy variable will unify all months to be the same, it will model some months properly but not all. On average, using a dummy might be a good idea but to fit the dummy to the model already estimated is a challenge. Using dummy variables can be useful and will be presented next. One will observe an estimated dummy model, which is visualized together with the actual sales in figure B in Appendix. The model consist only dummy variables and a constant. The dummy variables indicate January, August and December. All other months have value of zero and naturally, only one month will have value of one in the dummy variable, e.g., dummy variable for January will have the value of one only in January.

From the figure B presented in Appendix, one will see just how well the model actually fits. There are only three deviations from the constant in the model and the model captures the sales behavior surprisingly well. However, at the same time is presented the issue already mentioned, not all drops and peaks are with the same magnitude. One will see that there are drops in sales that are not captured by the model and then there are drops that did not happen, i.e., steady sales instead of a drop. Similar case is with the peaks, there are higher peaks and then there are peaks that are not in the scope of dummy variables, i.e., in different months than typically those will happen. With this demonstration, one will understand that the usage of simple dummy variables is not sufficient. If the dummy variable do not work always, it will cause in some cases more damage to the forecast than it actually brings value to those months were there are issues with the forecast accuracy. From business perspective, completely wrong forecast is worse than a model that is slightly but more often wrong.

Even though the dummies did not bring graphically satisfactory results, the descriptive results for the model are impressive. From table C presented in Appendix, one will see that the variable count adjusted $\bar{R}^2$ is surprisingly high at 39 %, which is in addition depressing. One can conclude that there has not been that many or any changes how the revenue recognition is done. From the table C presented in Appendix, one will also observe that all three variables have statistically significant coefficients, i.e., the t-probs are below 0.05. Naturally, the coefficients are exactly as one will expect those to be, DummyJAN and DummyAUG have negative coefficient whereas the DummyDEC has a positive coefficient. The F-test result is also statistically significant on the behalf the alternative hypothesis that the coefficients are not zero at the same time. All the residual tests are also passed, although the autocorrelation test just barely. The dummy variable analysis presented in the table C in Appendix reveal that the issue of odd sales behavior should be further evaluated. The usage of the dummy variables in the original model is evaluated next.

When all three dummy variables are added to the original model, the dummy variables are not anymore statistically significant. This can be seen from the table D presented in Appendix when observing the t-values for variables. One reason for this is due to the original model, which already captures the fundamental behavior of these months, but not all of it. The dummy variables do not add value to the model if one wants to have the coefficients of the dummy variables statistically significant. In addition, one will, because in this case the DummyJAN will have a positive coefficient, the opposite to what is expected, and was proven by the dummy variable model. If the DummyJAN is removed, then the DummyAUG will be close to the 5 % significance level at 5.2 %. At the same time, the DummyDEC is not statistically significant and will be removed next. However, then the DummyAUG will not be statistically significant and even when DummyDEC is added to model without the other two, it will not be statistically significant either. Neither will the DummyJAN be statistically significant when it is only dummy variable in the model. The exiting idea to improve the model with dummy variables is much harder to execute than initially thought.

Next, the idea of dummy variables is tested with variables Dummy-year and Dummy-year2, which are then compared against the original model. The Dummy-year has a value of one for January, February, November and December to highlight the effect of change in accounting period. Sales are boosted before the year-end and this will decrease the sales for next two months, i.e., during January and February. Dummy-year2 on the other hand is not a traditional dummy variable, but variable with changing magnitude to highlight the month specific issue. For January and February, the Dummy-year2 will get value of minus one. For November and December the opposite, i.e., one but for August the value is minus two. Fundamental issue with Dummy-year is the way the year-end months are now highlighted with the same signed value but the effect should be different depending on which side of the year-end one is. However, when November and December have different value than January and February, the variable is not anymore statistically significant. Due to this the analysis are done with this setup.

From tables E and F, which are presented in Appendix, one will see that all variables are statistically significant and the variable adjusted $\bar{R}^2$ is 59 % for the model with Dummy-year (table E in Appendix) and 61% respectively for the Dummy-year2 version (table F in Appendix). All residual test are fine and passed and the F-test is passed as well. The interest is with the forecast accuracy. Key statistics already used are the MAPE and RMSE. To test the forecast accuracy, the model must be estimated up to 2016 and then forecast performed for 2017. For the Dummy-year model, the RMSE is 0.19826 and the MAPE is 1.9249. Similar statistics for the Dummy-year2 model are RMSE is 0.24029 and MAPE 2.2595. The latter finding is crucial, even though the R^2 and adjusted $\bar{R}^2$ are higher in the latter model than for the former model, RMSE and MAPE are actually higher. On paper, the model with variable Dummy-year2 appeared to outperform the other models but in reality, it did not. Moreover, when one remembers how well the original model predicted the yearly sales for 2017, the latter model is not an upgrade but a downgrade in terms of forecast accuracy.

To benchmark against the original model, the adjusted $\bar{R}^2$ is 58% and the key statistics are RMSE was 0.20473 and MAPE was 1.9217. The fit of the model for the historical data can be reviewed with RSS value, for the original model it is 4.1041 and for the dummy models 3.8831 (Dummy-year) and 3.8056 (Dummy-year2) respectively. From this point of view, one would prefer the Dummy-year2 variable model to the other options but not based on the forecast accuracy for 2017.

Should one then move to the model, which uses the Dummy-year variable? Well it depends. One should be skeptical since the values of the dummy variable itself are not exactly what to expect. The dummy variable will force the model to re-estimate the regression coefficients and adjust the constant as well. All of these are appreciated features but still raises question, should this be done differently. Should one approach the non-fundamental behavior of sales internally, i.e., should one find internal variable that explains the effect of the affected months better? Since the monthly effect is not constant, one should have a variable with changing value and magnitude to indicate the issue in that particular month to adjust the forecast. For example, one could use invoicing activity versus ready-to-deliver machines, i.e., the sum of near future deliveries. Another variable could be deliveries of machines in the next month or the incapability to deliver machines in the previous month. With these additions, one could have a model that is both statistically significant and which is adjusted to capture the cyclic behavior of sales.

Study has so far concentrated to only one region out of two. The next step would be to analyze the APAC market and then create the model to that market as well. The issue rises how to get the data and what adjustments have to be made. The research started with all products and monitoring the different possibilities. US market is large for terminal tractors and was tempting but the sales process is not ideal. During the initial investigations, it came clear that from the US economy one can easily gather information. The Federal Reserve has the best coverage for financial data in the world; one will find unbelievably high coverage of industries and financial figures from one place and can export those using their Excel add-on. (Baumohl 2008, xx.) Moving to Europe means that the data will be in two places, some data would be at Eurostat and most of the financial data at the European Central Bank. Even this is manageable but to have overall level aggregate data sets from area, where there is no single governing body or coalition will cause difficulties. Most likely one needs to create similar indices by themselves and gather information from different sources. From this perspective, forklift truck modelling in APAC will be challenging but possible if the variables can be found or created.

Forklift trucks are only one of three profit centers in Kalmar Mobile Equipment. To have full coverage of Mobile Equipment, one should model the other profit centers as well. Container handling machines and terminal tractors both require different approach and fundaments than the forklift trucks. One should model the container handlers with container business and traffic and together with overall consumption, both businesses and customers. Terminal tractors are similar to container handlers but the input variable is not always seaside traffic but also the traffic with-in a country, i.e., produced, transported and consumed with-in one country. The tractors will need its own fundamental based model as well.

With preliminary studies, the container handling machines could be modelled with medium ease, which cannot be said about the terminal tractors. Sales of the container handlers behave rationally with a similar cycle, as does the container traffic. The terminal tractors are a complete opposite, the sales of tractors is highly cyclical which causes the issues that the business fundaments do not determine the sales value per period but customer needs, human bias and incentives do. The terminal tractors will need a model that gathers information from different sources, external fundaments and some internal indicators when the invoicing and deliveries will take place.

One of the first question that rose when the model was presented: “Could we have this for 12 months?” Naturally, this is healthy greed but has a valuable point. The model itself is not perfect to be used for static long-term forecasting but is usable for rolling forecasting. Only way one can extend it is to have a forecast of the underlying variables. Other option would be to create a separate model using for example quarterly data to forecast the higher aggregate level progress forward. In order to progress with this, one will need longer data set, since seven years data will not be long enough, it will be only 28 data points. One will quickly realize the issue with quarterly data; one will need 25 years of data to have a large sample. The company has gone through multiple changes in that time. From this perspective, creating the model on monthly level and then aggregating to quarterly level will be the most plausible way to progress. However, still one needs to have a solution to the forecast, it will not be possible to forecast fact based with this model any longer than the lags will enable. Most likely solution is to create a separate forecast model for the underlying variables. One remark must still be made. When fundaments drive the sales in a way discovered in this study, wish for a forecast model with longer forecast horizon do not make sense. In such case, there should be a separate study for about the underlying fundaments and forecasting practice for those.

The original inspiration and question for the research was could one find a causal relationship between macroeconomic variables and the sales and whether one can forecast sales using that relationship. Additionally the interest was the forecast horizon. Based on the research done in this thesis, forklift truck sales in EMEA region can be forecasted 11 months forward with macroeconomic variables by utilizing an OLS regression model. The model estimated is statistically sound and outperforms operations’ budget and do track the sales behavior accurately. On yearly level, the forecast model is highly accurate but on monthly basis, the forecast model can suffer from incompatibility with the revenue recognition process and customer deliveries.

The financial profitability the question, whether equipment is purchased based on financial market environment or because the machine can be fitted to budget, will remain to some extent open. The hockey stick effect (Chen 2000, 186) will imply the presence of purchasing equipment to fill the cost base in order to secure high cost budget for the next year as well. The hockey stick effect is not only about the customer, there are internal incentives to deliver as high results as possible. The profitability part of question was not covered in the research since the business fundamentals do explain accurately the historical data. Financial market information should have affect to sales but there are two possibilities. Firstly, the equipment can be purchased now because it will be more profitable. This was not proven in the study because it was overruled by the fundamental facts from the industries where the forklift trucks are operated. Secondly financial market environment will affect the industries where one operates with forklift trucks. In those industries, there will be more profitable investment possibilities to be performed when interest rates are low and due to the increase in production in the industries, the forklift truck sales will increase as well. Financial markets will have an effect but it is not modeled in the research since the effect to industries overruled the effect experienced in the equipment sales.

Based on previous research from the field, the econometric model will give an objective base line for budgeting and forecasting. The objectivity is the main attribute and value that the model will bring in normal cases since it will not have any incentives included, it will present facts as those are. Experts can have high forecast accuracy when there are no incentives or management desires included to adjust their forecast that they can actually produce. Once again, the human is the strongest and the weakest link in the forecasting profession. One can forecast accurately but may not be satisfied with the results and starts to manipulate the forecast to a better direction for their personal wellbeing. (Chase, 2013; Gilliland et al., 2015; Makridakis et al., 1998; Walker & McClelland 1991)

During the research, it came clear that the odd sales behavior is difficult to model and estimate. In the later part of the research, there was a dummy variable exercise to model the uncaptured behavior of sales. With dummy variables and even with changing magnitude it was not possible to forecast more accurately than with the original model. The dummy variable analysis raised the question for the future: how the customer behavior and internal invoicing activities can be modeled in time series environment. Can one find a sufficient times series or another indicator for it or should there be manual adjustment for the forecasts based the actual development in the previous months?

Based on experimenting and testing the model for two previous years independently and together, the forecast model will provide a clear indicator for monthly sales. The effect of revenue recognition between months is something the model do not capture accurately. Despite this, the original goal for the forecast model is met; the model does indicate the development of sales accurately when there are no internal issues or customer demands that are not normal. The goal for the research is met and it is possible to forecast using only fundamental facts as explanatory variables.

References

Literature

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, *91*(434), 444-455.

Armstrong, J. S. (1978). Forecasting with econometric methods: Folklore versus fact. *Journal of Business*, 549-564.

Baumohl, B. (2012). *The secrets of economic indicators: hidden clues to future economic trends and investment opportunities.* FT Press.

Brealey, Richard A. *Principles of Corporate Finance*. 11. ed., global ed. New York, NY: McGraw-Hill Education, 2014.

Bunn, D., & Wright, G. (1991). Interaction of judgemental and statistical forecasting methods: issues & analysis. *Management science*, 37(5), 501-518.

Cerullo, M. J., & Avila, A. (1975). Sales forecasting practices: A survey. *Managerial Planning*, 24(2), 33-39.

Chase, C. (2013). *Demand-driven forecasting: A structured approach to forecasting* (2nd ed.). Hoboken, N.J.: Wiley.

Chen, F. (2000). Sales-force incentives and inventory management. *Manufacturing & Service Operations Management*, 2(2), 186-202.

Corker, R. J., Holly, S., & Ellis, R. G. (1986). Uncertainty and forecast precision. *International journal of forecasting*, 2(1), 53-69.

D'Agostino, R. B. (1970). Transformation to normality of the null distribution of g1. *Biometrika*, 679-681.

Dawes, R. M. (1975). Graduate admission variables and future success. *Science*, 187(4178), 721-723.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427-431.

Donihue, M. R. (1993). Evaluating the role judgment plays in forecast accuracy. *Journal of Forecasting*, 12(2), 81-92.

Doomik, J. A., & Hansen, H. (1994). A practical test for univariate and multivariate normality. *Discussion paper*, *Nuffield College*.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: *Journal of the Econometric Society*, 987-1007.

Engle, R., & Granger, C. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica, 55*(2), 251-276. doi:10.2307/1913236

Gilliland, M., Sglavo, U. & Tashman, L. (2015). *Business forecasting: Practical problems and solutions*. Hoboken, New Jersey: Wiley.

Gourieroux, C. & Jasiak, J. (2001). *Financial econometrics: Problems, models, and methods*. Princeton: Princeton University Press.

Granger, C. W. (1981). Some properties of time series data and their use in econometric model specification. *Journal of econometrics*, *16*(1), 121-130.

Granger, C. W. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of economics and statistics*, *48*(3), 213-228.

Greene, W. H. (2012). *Econometric analysis* (7th ed.). Harlow : Pearson.

Harvey, A. C. (1990). *The econometric analysis of time series*. Mit Press.

Hanke, J.E. & Wichern, D.W., 2009. *Business forecasting*. Upper Saddle River, NJ: Pearson, 2009.

Heij, C. (2004). *Econometric methods with applications in business and economics*. Oxford: Oxford University Press.

Klein, H. E., & Linneman, R. E. (1984). Environmental assessment: An international study of corporate practice. *Journal of Business Strategy*, 5(1), 66-75.

Maddala, G. S. & Kim, I. (1998). *Unit roots, cointegration and structural change*. Cambridge: Cambridge University Press.

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting methods and applications*. John wiley & sons.

Mathews, B. P., & Diamantopoulos, A. (1986). Managerial intervention in forecasting. An empirical investigation of forecast manipulation. *International Journal of Research in Marketing*, 3(1), 3-10.

McNees, S. K., & Perna, N. S. (1981). Forecasting macroeconomic variables: An eclectic approach. *New England Economic Journal*, 15-30.

Meehl, P. E. (1954). Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.

Mendenhall, W., & Reinmuth, J. E. (1993). *Statistics for management and economics*.

Morlidge, Steve, and Steve Player. *Future Ready: How to Master Business Forecasting*. Hoboken, N.J.: Wiley, 2010.

Nahmias, S., 2009. *Production and operations analysis*. New York, NY: McGraw-Hill/Irwin, cop. 2009.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society*. Series B (Methodological), 350-371.

Reinmuth, J. E., & Geurts, M. D. (1976). A Bayesian approach to forecasting the effects of atypical situations. *In Readings in Managerial Economics* (pp. 137-147).

Rothe, J. T. (1978). Effectiveness of sales forecasting methods. *Industrial Marketing Management*, 7(2), 114-118.

Samuelson, P. A. (1973). Proof that properly discounted present values of assets vibrate randomly. *The Bell Journal of Economics and Management Science*, 369-374.

Shenton, L. R., & Bowman, K. O. (1977). A bivariate model for the distribution of√ b1 and b2. *Journal of the American Statistical Association*, 72(357), 206-211.

Shim, J. K., Siegel, J. G. & Shim, A. I. (2012). *Budgeting basics and beyond* (4th ed.). Hoboken, N.J.: Wiley.

Silver, E. A., Pyke, D. F., & Peterson, R. (1998). *Inventory management and production planning and scheduling* (Vol. 3, p. 30). New York: Wiley.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3), 165-180.

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management science*, 36(6), 643-660.

Turner, D. S. (1990). The role of judgement in macroeconomic forecasting. *Journal of Forecasting*, 9(4), 315-345.

Verbeek, M. (2008). *A guide to modern econometrics*. John Wiley & Sons.

Walker, K. B., & McClelland, L. A. (1991). Management forecasts and statistical prediction model forecasts in corporate budgeting. *Journal of Accounting Research*, 371-381.

Webby, R., & O'Connor, M. (1996). Judgemental and statistical time series forecasting: a review of the literature. *International Journal of Forecasting*, 12(1), 91-118.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica: *Journal of the Econometric Society*, 817-838.

Web resources

Cargotec (2015) CEO review at Cargotec's Annual General Meeting: Cargotec's 2016 guidance unchanged (accessed 28.11.2017) https://www.cargotec.com/en/nasdaq/stock-exchange-release-kalmar-hiab-macgregor/2016/ceo-review-at-cargotecs-annual-general-meeting-cargotecs-2016-guidance-unchanged/?_t_id=1B2M2Y8AsgTpgAmY7PhCfg%3d%3d&_t_q=guidance+2016&_t_tags=language%3aen&_t_ip=134.37.0.9&_t_hit.id=Pure_Core_Models_Pages_ContentPage/_0bf22388-8412-4d02-818b-579f6e9d9bc2_en&_t_hit.pos=1

Cargotec (2010) Financial Statements Review 2009 (accessed 28.11.2017)

https://www.cargotec.com/globalassets/files/investors/presentations/q4-2009-presentation-en-final.pdf

Chockalingam, M. (2009). What is true demand? Demand planning Newsletter (April) (accessed 28.11.2017) http://www.dem&planning.net/newsletters/DPnewsletter_april2009.html

Kalmar (2017) Kalmar unveils again an industry first: Fully electric medium-range forklift (accessed 26.2.2018) https://www.kalmarglobal.com/pressroom/press_releases/2017/kalmar-unveils-again-an-industry-first-fully-electric-medium-range-forklift-/

Konecranes (2017) Konecranes Interim report Q1 / 2017 (accessed 28.11.2017) http://www.konecranes.com/sites/default/files/investor/kc_2017_q1_en.pdf

Metso (2017) Metso issued a EUR 300 million Eurobond https://www.metso.com/news/2017/5/metso-issued-a-eur-300-million-eurobond/

Nokian Tyres (2018) Review by the President and CEO (accessed 26.2.2018) https://www.nokiantyres.com/company/investors/nokian-tyres-as-an-investment/review-by-the-president-and-ceo/

Valmet (2017) Valmet revises upwards its net sales guidance for 2017 (accessed 28.11.2017) http://www.valmet.com/media/news/stock-exchange-releases/2017/valmet-revises-upwards-its-net-sales-guidance-for-2017/

Other references

Institute of Shipping Economics and Logistics (2017) - ISL container throughput CMM201707 - Consultancy data set

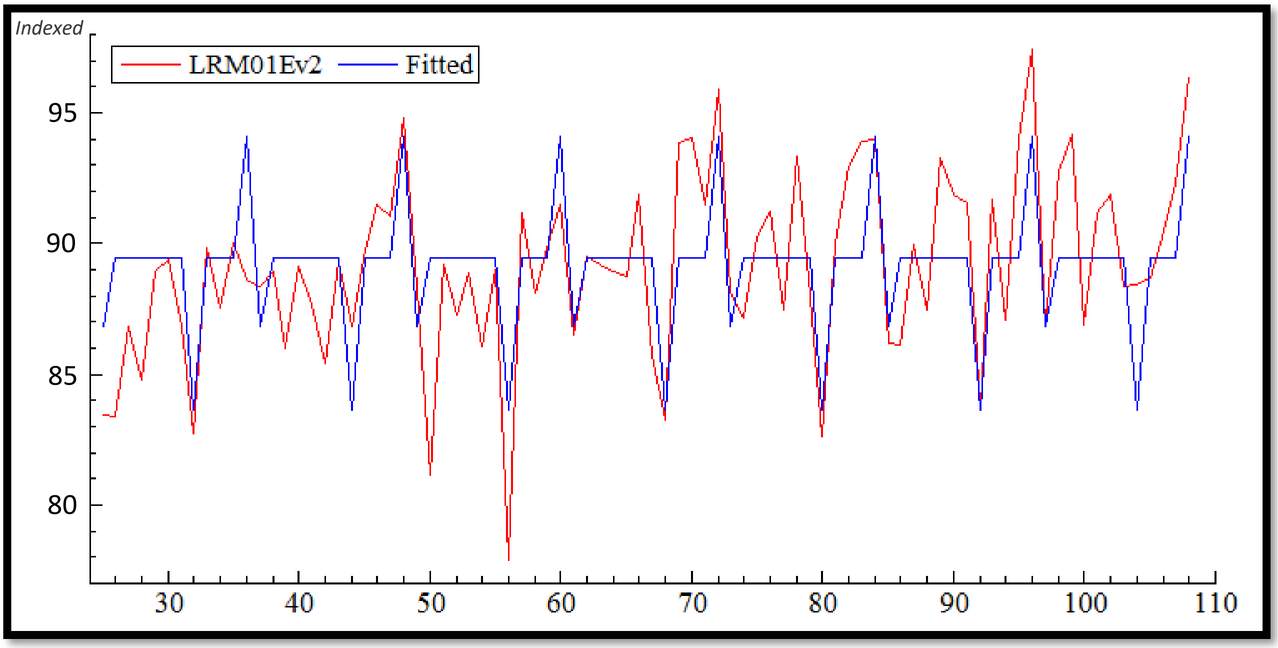
Appendix

Appendix A. Mean and standard error for the forecasted periods (indexed to actual January, level 100)

Horizon	Forecast	SE	Actual	Error	t-value	-2SE	+2SE
January	99.87	100.00	100.00	0.13	0,044	94.10	105.65
February	103.72	102.33	99.96	-1.39	-1,272	97.81	109.62
March	104.50	97.83	104.38	-6.67	-0,042	98.85	110.15
April	104.17	99.24	101.48	-4.93	-0,939	98.44	109.90
May	107.8	105.18	108.25	-2.62	0,149	101.73	113.87
June	106.68	100.52	106.64	-6.16	-0,016	100.88	112.49
July	102.94	99.32	106.23	-3.62	1,145	97.21	108.68
August	100.99	100.04	97.20	-0.95	-1,31	95.21	106.77
September	105.60	101.13	106.40	-4.47	0,274	99.76	111.44
October	105.75	99.52	101.04	-6.23	-1,638	100.00	111.49
November	106.32	100.16	109.05	-6.16	0,945	100.53	112.10
December	109.83	102.77	113.06	-7.06	1,089	103.89	115.77
January	101.39	99.40	100.76	-1.99	-0,219	95.65	107.13
February	105.37	103.34	107.69	-2.03	0,777	99.49	111.34
March	105.86	99.64	109.25	-6.22	1,18	100.11	111.61
April	105.51	98.79	100.83	-6.72	-1,64	99.80	111.21
May	107.38	103.66	105.82	-3.72	-0,523	101.39	113.37
June	105.66	100.24	106.62	-5.42	0,333	99.87	111.45
July	104.84	99.52	102.55	-5.32	-0,799	99.10	110.59
August	101.74	101.05	102.63	-0.69	0,306	95.90	107.57
September	105.13	101.13	102.94	-4.00	-0,75	99.29	110.97
October	108.22	102.61	105.00	-5.61	-1,088	102.30	114.15
November	107.16	101.00	107.12	-6.16	-0,014	101.33	112.99
December	110.41	103.94	111.81	-6.47	0,466	104.41	116.42

Estimated by the author.

Appendix B. Model fit between dummy variable model (fitted) and actual sales (LRM01Ev2).



Estimated by the author.

Appendix C. Econometric model estimation using Dummy variables for January, August and December. Sales is the dependable variable.

```
                  Coefficient  Std.Error  t-value  t-prob Part.R^2
Constant              89.4388    0.03476            0.0000   0.9988
DummyJAN            -0.261200     0.1099    -2.38   0.0199   0.0659
DummyAUG            -0.578609     0.1099    -5.26   0.0000   0.2573
DummyDEC             0.464043     0.1099     4.22   0.0001   0.1822

sigma                0.275872  RSS                 6.08841867
R^2                  0.410865  F(3,80) =      18.6 [0.000]**
Adj.R^2              0.388772  log-likelihood        -8.96484
no. of observations        84  no. of parameters            4
mean(LRM01Ev2)        8.91257  se(LRM01Ev2)          0.352863

AR 1-2 test:      F(2,78)   =   2.8121 [0.0662]
ARCH 1-1 test:    F(1,82)   =  0.97676 [0.3259]
Normality test:   Chi^2(2)  =   2.3731 [0.3053]
Hetero test: no regressors for test
Hetero-X test: no regressors for test
RESET23 test:     F(2,78)   =3.5204e-028 [1.0000]
```

Estimated by the author.

Appendix D. Original econometric model with dummy variables for January, August and December. Sales is the dependable variable.

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-34.3638	9.791	-3.51	0.0008	0.1411
Calendar ManufacturingUpdate_11	0.0264632	0.008676	3.05	0.0032	0.1104
Calendar ManufacturingUpdate_15	-0.0137000	0.005216	-2.63	0.0105	0.0842
Calendar ConstructionUpdateV2_11	-0.0245410	0.008688	-2.82	0.0061	0.0962
Calendar ConstructionUpdateV2_12	0.0257236	0.005918	4.35	0.0000	0.2012
LEU28 MOVEMENT QUANTITY_IN_100KG/1000_14	2.89297	0.6909	4.19	0.0001	0.1895
DummyJAN	0.139051	0.1286	1.08	0.2831	0.0153
DummyAUG	-0.182645	0.1265	-1.44	0.1529	0.0271
DummyDEC	0.217315	0.1378	1.58	0.1189	0.0321

sigma	0.225558	RSS	3.81571712
R^2	0.630779	F(8,75) =	16.02 [0.000]**
Adj.R^2	0.591395	log-likelihood	10.6601
no. of observations	84	no. of parameters	9
mean(LRM01Ev2)	8.91257	se(LRM01Ev2)	0.352863

AR 1-2 test:	F(2,73)	=	0.59270 [0.5555]
ARCH 1-1 test:	F(1,82)	=	1.1509 [0.2865]
Normality test:	Chi^2(2)	=	0.64096 [0.7258]
Hetero test:	F(13,70)	=	1.3093 [0.2286]
Hetero-X test:	F(23,60)	=	1.4714 [0.1175]
RESET23 test:	F(2,73)	=	0.049551 [0.9517]

Estimated by the author.

Appendix E. Original econometric model with Dummy-year variable included. Sales is the dependable variable.

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-32.5475	9.737	-3.34	0.0013	0.1267
Calendar ManufacturingUpdate_11	0.0287251	0.007402	3.88	0.0002	0.1636
Calendar ManufacturingUpdate_15	-0.0117410	0.004764	-2.46	0.0160	0.0731
Calendar ConstructionUpdateV2_11	-0.0308626	0.006293	-4.90	0.0000	0.2380
Calendar ConstructionUpdateV2_12	0.0310712	0.003545	8.77	0.0000	0.4995
LEU28 MOVEMENT QUANTITY_IN_100KG/1000_14	2.74146	0.6893	3.98	0.0002	0.1704
Dummy-year	0.128420	0.06133	2.09	0.0396	0.0539

sigma	0.224565	RSS	3.88306616
R^2	0.624262	F(6,77) =	21.32 [0.000]**
Adj.R^2	0.594984	log-likelihood	9.92522
no. of observations	84	no. of parameters	7
mean(LRM01Ev2)	8.91257	se(LRM01Ev2)	0.352863

AR 1-2 test:	F(2,75)	=	0.61435 [0.5437]
ARCH 1-1 test:	F(1,82)	=	1.2920 [0.2590]
Normality test:	Chi^2(2)	=	1.0774 [0.5835]
Hetero test:	F(11,72)	=	1.1108 [0.3657]
Hetero-X test:	F(21,62)	=	1.2366 [0.2547]
RESET23 test:	F(2,75)	=	0.40229 [0.6702]

Estimated by the author.

Appendix F. Econometric model using 11 month lags for construction, manufacturing, and movement variable together with Dummy-year2 variable. Sales is the dependable variable.

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-24.6487	9.394	-2.62	0.0104	0.0802
Calendar ManufacturingUpdate_11	0.0162129	0.005661	2.86	0.0054	0.0941
Calendar ConstructionUpdateV2_11	-0.00868532	0.004058	-2.14	0.0354	0.0548
LEU28 MOVEMENT QUANTITY_IN_100KG/1000_14	2.26653	0.6567	3.45	0.0009	0.1310
Dummy-year2	0.319906	0.03460	9.25	0.0000	0.5198

sigma	0.219484	RSS	3.80568507
R^2	0.63175	F(4,79) =	33.88 [0.000]**
Adj.R^2	0.613104	log-likelihood	10.7706
no. of observations	84	no. of parameters	5
mean(LRM01Ev2)	8.91257	se(LRM01Ev2)	0.352863

AR 1-2 test:	F(2,77)	=	0.48896 [0.6152]
ARCH 1-1 test:	F(1,82)	=	1.4279 [0.2356]
Normality test:	Chi^2(2)	=	0.67515 [0.7135]
Hetero test:	F(8,75)	=	1.7373 [0.1036]
Hetero-X test:	F(14,69)	=	1.3078 [0.2257]
RESET23 test:	F(2,77)	=	0.31545 [0.7304]

Estimated by the author.