UNIVERSITY OF TURKU

ABSTRACT

	Bachelor's thesis
X	Master's thesis
	Licentiate's thesis
	Doctor's thesis

Subject	Economics	Date	16.12.2020
Author	Oskari Määttä	Number of pages	54
Title	Forecasting U.S. Industrial Production - Does the Baltic Dry Index provide additional predicting power		
Supervisor	Prof. Heikki Kauppi		

Abstract

Various economic leading indicators are continuously referred in the financial discussion. Asset managers, central bankers and trading participants pay attention to myriads leading indicators and base their decisions on the information of the expected economic performance. This study investigates the relationship between leading indices and US industrial output and aims to provide further knowledge by evaluating the forecast accuracy of leading indices. Furthermore, the capabilities of Baltic Dry Index as a leading indicator is evaluated alongside with asserting momentum's usefulness to enhance forecasting accuracy.

Recursive expanding estimation window forecasts were created for US industrial production. Forecasts are based on the on the estimated time series of Conference Boards Leading Index, Baltic Dry Index, US government bond yield spread and the autocorrelation of the of the dependent variable. Robust empirical analysis concluded that Baltic Dry Index does not have predictive nature for US industrial production. In addition, momentum enhances the forecasting performance of Conference Board Leading Index when autocorrelation variable was included in the forecasting model with the Leading Index to predict the US industrial output.

These results suggest that Baltic Dry Index should not be considered as a leading indicator and Conference Boards Leading Index should not include it as a constituent. However, momentum in a for of autocorrelation should be considered within the Conference Boards Leading Index as it enhances its forecasting accuracy.

Key words	Leading Indicators, Forecasting

TURUN YLIOPISTO

Tiivistelmä

	Kandidaatintutkielma
X	Pro gradu -tutkielma
	Lisensiaatintutkielma
	Väitöskirja

Oppiaine	Taloustiede	Päivämäärä	16.12.2020
Tekijä	Oskari Määttä	Sivumäärä	54
Title	Yhdysvaltojen teollisuustuotannon ennustaminen – Lisääkö Baltic Dry Indeksi ennustevoimaa		
Ohjaaja	Prof. Heikki Kauppi		

Tiivistelmä

Taloudellisia indikaattoreita seurataan taloudellisessa keskustelussa jatkuvasti. Varainhoitajat, keskuspankkiirit ja kaupankäyntiosapuolet huomioivat taloudellisten indikaattorien tuomaa informaatiota tehdäkseen päätöksiä tulevaisuuden todennäköisen taloudellisen tilan perusteella. Tämä tutkimus käsittelee Yhdysvaltojen teollisuustuotannon ja johtavien taloudellisten indikaattorien välistä suhdetta tarkoituksenaan löytää informaatiota indikaattorien selitysvoimasta. Lisäksi arvioidaan Baltic Dry Indexin käyttökelpoisuutta johtavana indikaattorina sekä momentumin hyödyntämistä johtavien indeksien ennusteissa.

Yhdysvaltojen teollisuustuotannolle luotiin rekursiivisia ennusteita, jotka hyödyntävät ekspansiivista estimointi-ikkunaa. Ennustemallien estimaatit perustuvat Conference Board Leading Indexiin, Baltic Dry Indexiin, Yhdysvaltojen valtiovelkakirjojen korkojen väliseen eroon sekä selitettävän muuttujan autokorrelaatioon. Robustin analyysin perusteella Baltic Dry Index ei sisällä ennustevoimaa Yhdysvaltojen teollisuustuotannolle. Lisäksi momentumin lisääminen ennustemalliin tarkensi ennustemalleja merkittävästi, kun Yhdysvaltojen teollisuustuotantoa selittävinä muuttujina olivat autokorrelaatio ja Conference Board Leading Index.

Näiden tulosten perusteella Baltic Dry Index ei ole johtava taloudellinen indikaattori ja sitä ei tulisi lisätä Conference Board Leading Index komponentiksi. Sen sijaan momentum tulisi huomioida Conference Board Leading Indexissä, koska sen lisääminen ennustemalliin tarkentaa ennustetta merkittävästi.

Key words	Johtavat indikaattorit, Ennustaminen

FORECASTING U.S. INDUSTRIAL PRODUCTION

Does the Baltic Dry Index provide additional predicting power

Master´s Thesis
in Economics

Author:
Oskari Määttä

Supervisor:
Prof. Heikki Kauppi

16.12.2020
Turku

Table of Contents

List of Figures

List of Tables

1 INTRODUCTION

Financial institution and economical decision makers use various indicators to predict and estimate the state of the business cycle to set inflation targets and policy rates. Predicting the turning point of the cycle has intrigued the central bankers and investors since the early stages of macroeconomic studies. The opening section will briefly introduce the motivation, research background, purpose and the structure of the thesis.

After years of following the everlasting financial discussion about the end of the bull market and the growth of the economy, I wanted to dive deeper into the topic of economic indicators and their level of relevance. At a time when the world economy has just crashed due Covid19 virus and the stock market it somehow still going strong, it is intriguing to look for indicators that could reveal a glimpse of the future. The leading indicators are followed intensively on a global scale from asset managers to economist. Scheduled economic issues published by central banks and index providers are long known to cause major intraday volatility on centrally traded assets (Ederington & Lee, 1993). However, their appropriateness to properly signal either the pace or the direction of economic growth has been criticized early on in the academia due lack of theoretical background (Auerbach, 1982). From a personal perspective it is interesting to gather the information of different types of leading indicators because the literature and the theory of the leading indicators seems in far away from cohesive. The purpose is to highlight the key theories behind the commonly referred indicators and to test if the indicators do in fact lead. Reason why Baltic Dry Index, and its reliability as a leading indicator, is taken into the consideration since its wide-spread public coverage in financial industry. Major financial news outlets refer to it regularly and new indices such as Index of Global Real Economic Activity has been built based on its values (Lutz, 2009).

Section 2 will initiate the reader to the previous research of leading indicators in a more detailed manner. In the subchapters, the reader is walked through the key research used as a background information when evaluating which time series are used later on. Perhaps the most influential research of leading indicators is related to the predictive power of the interest rates. The theory of the business cycles and interested rate curve is heavily influenced by the work of the American economist Frederic S. Mishkin. This thesis introduces the basic concepts of the yield curve and refers several times to Mishkins papers from the late 1980s. His papers from 1988 and 1989 seem to have created a building block to further studies since they have also been used in various other sources referred in this study. Second chapter also briefly introduces the reader to the research of using other financial instruments than yield curve as a leading indicator such as liquidity premiums in equity and repo markets. Furthermore, the leading indices are covered where multiple indicators are compiled to a single index to provide general outlook of the overall expectation of the economy. The first research question is if leading indicators are as essential as they claim to be since the previous research either rejects their reliability (Diebold & Rudebusch, 1991) or favors them (Heij, Dijk, & Groenen, 2009). Autoregressive models are used as a minimum criterion and the models using external inputs should presumably outperform them to qualify as informative.

Section 3 introduces the reader to Baltic Dry Index to have a closer look at the real economic activity through global shipping prices. Third chapter familiarizes what economic forces drive global shipping prices. Learning more about the factors affecting to the shipping rates is relevant to evaluate whether Baltic Dry Index signals movement in economic activity. To gain further perspective, the methodology of the Baltic Dry Index is decomposed, and the reader is acquainted with the key research findings of the shipping index. The second research question relates to Baltic Dry Index. If leading indicators are considered as reliable, should Baltic Dry Index be used as one of the components?

Instead of just going through the past research of already applied financial and non-financial economic indicators, a new perspective is added to the framework by applying a new component to already tried empirical method. Intention is to find out if not only leading indicators can be practically useful but also to compare their forecasting power between each other. Fourth chapter describes the methodology of the empirical part. I will present the time series used in the regression model and cover the theory behind the statistical tests used to evaluate the validity of the data. Section 4 will introduce the reader to univariate, multivariate and combined approach of estimating and forecasting US Industrial Production. Combined approach refers to a linear estimation model where autoregressive construction is used alongside with external independent variable. This is to find out whether the leading indicators could benefit from using momentum as a parameter. In addition, evaluation criteria for each model are defined. Forecasts follow simulation out of sample framework where each estimation is fitted with the data available at the time of the event. The forecasted values are then compared to the actual values. Estimation window for each model is from January 2010 to December 2018 and a expanding estimation window approach is enabled where each forecasting model is fitted latest available observation.

Section 5 presents the empirical study and the results. Univariate model constructions are represented, and the lag structure is selected based on the autocorrelation parameters. Estimation outputs are covered briefly and presented at a single point of time. The fifth chapter focuses on the outcome of the recursive expanding estimation window forecasts. Ten one-step-ahead forecasts are generated for US industrial production where forecast horizon corresponds with the lag term used in the independent variables. Performance of each forecast is evaluated against each other and measured with appropriate metrics. Method of using 9-year estimation window with monthly values, ensures that the results are not just interpretations of a single random event but robust and applicable for conclusive analysis.

The final section will conclude the thesis by emphasizing the empirical results and their standing within the research framework of leading indicators. The validity of the study is evaluated against the previous research discoveries. Additionally, possible future research topics based on the empirical finding of the study are briefly discussed in the end of section 6.

2 ECONOMIC INDICATORS

The modern-day theory of economic indicators derives from the assumption that the actual economic output fluctuates around its potential output. However, the initial rationale behind the matter was simple, trying to get information whether economy is about to turn better and estimate the timing of the turning point. Mitchell and Burns (1938) tried to estimate the turning points of economic revival. They evaluated 487 indices provided by US statistical department and were able to extract 71 relevant to the economic cycle. As early as in this study the idea of leading and lagging indicators was present and the indices relevant to the business cycle were categorized to be either lagging or leading.

As the macroeconomic science developed further, the applications of leading indicators became more advanced. Not only was it desirable to estimate the turning point from the recession, but also to use indicators continuously as a tool to assess the current state of the economy, regardless of the current state of the economy. Forecasting and predicting recessions became an objective for the economist. This was nothing new in way since W.M Persons (1919) already introduced his idea of Economic cycles following a pattern. These patterns became known as Harvard-ABC Curves. However the approach at the time was arguably over simplified since economist failed to forecast the Great Depression of 1930s, post-World War II expansion of economy, the recessions of 1949 and 1957 in addition to inability to recognize the inflation environment of mid-1960s (Moore, 1983). Neither academic research nor policymaking seemed to be particularly successful regarding leading indices during the post war decades, but topic became increasingly popular later on. One reason for the resurgence of “measurement without theory approach” (Auerbach, 1982) was the progress of new more powerful econometric research tools. This enabled economists to apply modern econometric methods to forecast the economy based on either single or multiple variables.

2.1 Treasury yield curve

US treasury yield spread is widely considered as a leading indicator. The yield spread described in this study consists of the difference of yields between two fixed income instruments: 2-year maturity US government issued bond and the 10-year maturity US government issued bond as stated in figure 1. In a normal economic state, the riskier fixed income assets are the ones having longer maturities. The risk related to government bonds and especially US treasury bonds is the variation in interest rate and inflation over time (Cochrane & Piazzesi, 2005). Logically existing bonds will lose value if the central bank increases interest rate, because equivalent bonds issued after the rate change now bear higher coupon. This simply means that 10-year US government bond should yield higher interest rate than a respective 2-year bond since the interest rate risk is higher and therefore more premium is required. Normally this is the case and therefore spread is usually positive and the yield curve of compounded maturities is upward sloping as shown in the figure 2.

As central banks determine the short-term rate then arguably the market judges their decision-making by pricing the long-term interest rate. Long-term interest rate should be more sensitive and that the position of the yield curve would therefore indicate the direction of the economy based on the expectations of the market itself. Flattening curve would hint recession ahead and steep yield curve would indicate that central bank and the government are likely to stimulate the economy by loosening the interest rate and therefore an upward cycle would be likely. (Estrella & Hardouvelis, 1991)

Figure 1 US 10-year 2-year Treasury Yield Spread (FRED databank)

The yield spread has turned negative before the past three major downturns in the economy including recessions of the early 1990s, 2000s and the great recession of 2008 (Figure 1). The inverted yield spread has consistently preceded a near term recession. In fact, with a quick glance the predictable power of the yield spread seems almost as a self-fulfilling prophecy. Nevertheless, the yield curve compiles a significant amount of information through capital markets in to single figure and the interpretations derived of its direction is a heavily researched topic. (Estrella & Mishkin, The Yield Curve as a Predictor of U.S. Recessions, 1996)

Using the yield curve, as an economic indicator has become an industry wide standard and the yield curve used as a component in wider economic indexes introduced later in this study. However, in the academia, economists have tried to explain the shape yield curve trough four major theories. It is important to understand the forces that the shape the yield curve in a single point of time, which then the yield spread reflects as a single figure changing over longer timeframe as shown in figure 1. These four theories are pure expectation theory, liquidity theory, market segmentation theory and preferred habitat theory.

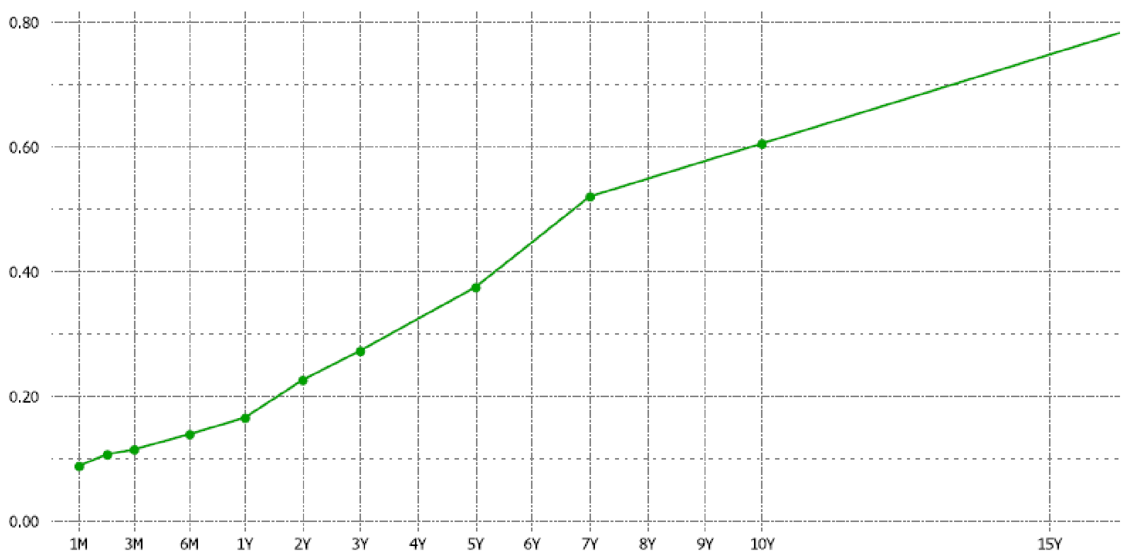


Figure 2 US Treasury upward yielding curve. (Bloomberg)

2.1.1 *Pure expectations theory*

According to the pure expectations theory of term structure, the long-term interest rate should be equal with weighted average of present and expected future short-term interest rates (Mankiv & Summers, 1984). Pure expectation theory also assumes that a normal concave yield curve, where longer maturities yield higher interest rate, would predict the future short-term interest rates to be on the rise. However, in a situation when the spread turns negative and the yield curve is inverted as shown in figure 3, it cannot hold that the same bond issuer would have less risk with longer maturity bond (Nymand-Andersen, 2018).

One way to reason the unusually small or even negative difference between short maturity yield and long maturity yield is to consider inflation expectations. Inflation is a key component in yield curve discussion. As Fischer effect (Fischer, 1930) states the real interest rate consists from both inflation and nominal interest rate. If the future inflation is expected to be lower than the current term inflation, the bond premium is less as well due the lesser inflation effect on profit (Mishkin F. S., 1988). This would lead to a situation where longer maturity bonds have lower yields than before and the yield spread would diminish (Mishkin F. S., 1989).

Naturally, the logic of the term structure works other way around as well. Meaning that the yield spread might be tightening due an increase in the short-term side of the equation. When central banks set their policy rates, it mostly affects to the near maturity yields. E.g. when Federal Reserve Open Market Committee decides to increase the federal funds rate, it would push the short-term interest rates up since more premium is now required due higher central bank rate.

Figure 3 US Treasury yielding curve inversion. Shorter maturities yielding higher returns. (Bloomberg)

2.1.2 *Liquidity preference theory*

Since the treasury yield curve is one of most followed indicators of economic cycle there are also other perspectives in the research than just pure expectations. The concept of liquidity has widely been associated with yield curves behavior. According to traditional asset pricing theory, every asset is valued based on the present value of future cash flows. This theory is however based on the assumption of frictionless market and the liquidity theory provides an alternative proposal.

Traditional Liquidity preference theory suggests that investors prefer liquidity available and therefore longer-term maturities trade with higher premium. Normal upward concave yield curve would then be explained by preference in cash. (Nymand-Andersen, 2018). Short-term premiums were already noticed by Mankiv and Lawrence (1984) and found out to be liquidity based especially on the very short end of the curve. Longstaff (2004) also noticed evidence for liquidity premiums even among equivalent short-term US bonds. Investors paid premiums for government issued bonds that had the same coupon and interest rate risk as the comparable government bond, but the liquidity pool was deeper. This behavior is defined as flight-to-liquidity.

2.1.3 *Behavioral perspective to yield curve*

In addition to expectations and liquidity, researchers have also considered market segments to play a role in yield curve's term structure. Market segmentation hypothesis states that term structure is defined by supply and demand of bonds in each maturity and therefore the shape of the yield curve if determined based on the number of investors and trading volume in the respective maturity (Johnson, Zuber, & Gandar, 2010). In other words, market segmentation theory implicitly assumes that each maturity has its own unique market which are unaffected by each other.

Preferred habitat theory adds on to this behavioral economics approach by suggesting that each type of investors prefer their habitat and therefore invest only on certain maturity bonds unless there is a premium or discount available on other maturity bonds. (Diebold & Rudebusch, Yield Curve Modeling and Forecasting, 2013). The foundation for the preferred habitat of certain maturity instruments is that investors and borrowers need to match the maturities of their assets and liabilities (Johnson, Zuber, & Gandar, 2010).

2.2 Financial market

Theoretically, the Stock market could work as a leading indicator for the economy measured. The background derives from the idea that each company's valuation is based on the present value of its future income. The challenge of this approach derives from the fact that often companies do have high valuation without having any or low projected income and the market valuates the company based on the high expectations of future performance. Also, the inconsistent empirical results of the stock prices predicting economic activity support the argument of non-rational behavior of the equity valuations (Stock & Watson, 2003).

There are other approaches than just price related indications of the stock market. Flight-to-Liquidity or Flight-to-Quality is a concept of liquidity indroduced by Francis A. Longstaff in his 2004 paper. His paper describes investors pulling out from riskier or less liquid assets and moving funds to assets considered less risky or more liquid assets in the U.S. Bond market. However, Naes, Skjeltorp and Odegaard (2011) applied the flight-to-liquidity approach to stock market. They found out that during economic uncertainty, investors tend to allocate their equity position away from less liquid small-cap or mid-cap shares towards more liquid blue chip companies subsequently boosting up their value due increased demand on the buy side. Furthermore, their study showed that the shift in liquidity from the small-cap and mid-cap shares to large-cap shares in fact happens before the coincident indicators, US real GDP in this case, show worsening economic performance. Flight-to-Liquidity would then qualify as a leading indicator as it seems to possess a predictive nature.

On top of concentrating to the liquidity of the share itself, it is also meaningful to consider the liquidity of the funding required to acquire financial instruments. In fact, the balance sheets of the financial institutions providing the liquidity show cyclical properties. Adrian and Hyung (2008) showed strong correlation between monetary policy and balance sheet growth of financial intermediaries. Their conclusion was that the repo market, and therefore the liquidity of the financial system itself, was highly cyclical and tightly related to central banks policy rate. According their study, instead of just signalling the market about the expected returns of future via the yield curve, the central bank policy rate actually drives the liquidity of the repo market directly. The growth of the repo market, due lower rates, inflates the balance sheets of the banks and eventually affects on the asset prices itself. Higher asset prices again enable more lending because the balance sheets are now stronger. The pro-cyclicality of leverage offers an interesting approach because it indirectly suggests that the stock market and financial assets in general are not forward looking and their could be actually artificially inflated by the central bank expanding its balance sheet extensively,

2.3 Leading Indices

Outside of financial markets there are also other ways to measure the state of the business cycle. One way is to combine various economic indicators to a wider index, which include indicators from financial markets to real economy. Aim of these indices is to measure and forecast the development of the business cycle.

Conference Board publishes composite indices, which combine multiple indicators together to form single indices trying to describe the state of the U.S. economy. Conference board publishes three composite indices: Leading Economic Index (LEI), Coincident Economic Index (CEI) and Lagging Economic Index (LAG) (The Conference Board, 2020). This study emphasizes attention to the leading indicators and therefore concentrating on Leading Economic Indicator (LEI) is justified. Conference Board Leading Index combines data from following indicators:

1 Average weekly hour, manufacturing
2 Average weekly initial claims for unemployment insurance
3 Manufacturers' new orders, consumer goods and materials
4 ISM® new orders index
5 Manufacturers' new orders, non-defense capital goods excl.
aircraft
6 Building permits, new private housing units
7 Stock prices, 500 common stocks
8 Leading Credit Index
9 Interest rate spread, 10-year Treasury bonds less federal funds rate
10 Average consumer expectations for business conditions

Each of the factors within the index are weighted based on the monthly change in the series $x_t = X_t - X_{t-1}$. However, to smooth out the volatility of individual series the individual contribution in the instrument is adjusted with the standardization factor. The adjusted contribution in each component is the monthly contribution multiplied by the corresponding component standardization factor ($m_t = r_x * x_t$). (The Conference Board, 2020).

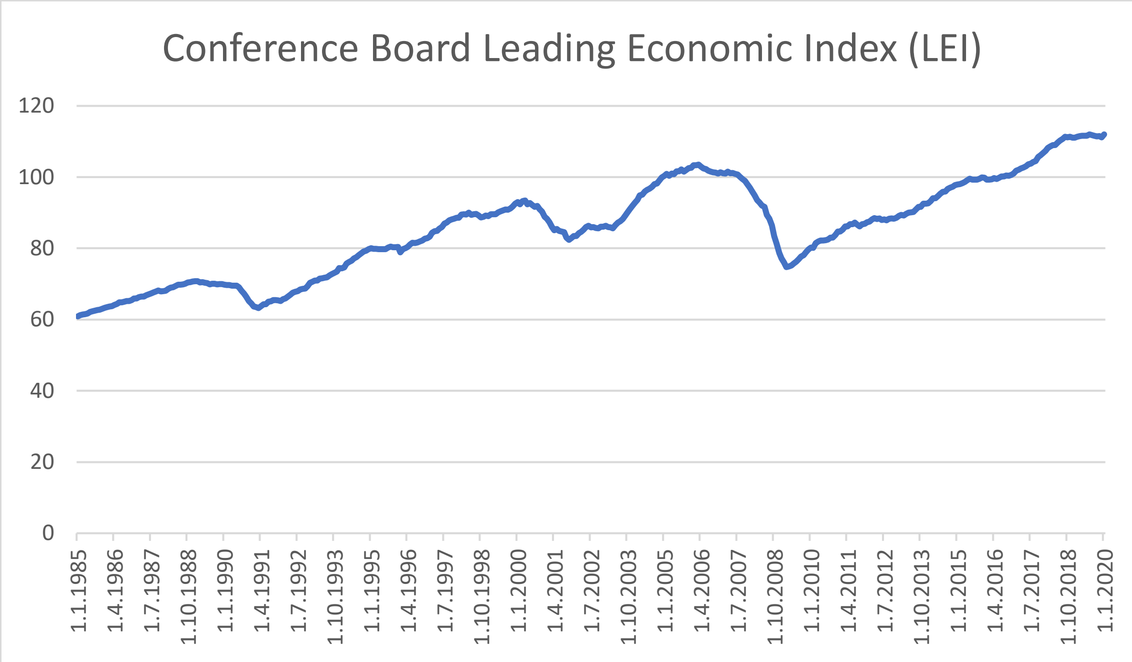


Figure 4 Leading Economic Index 1985-2020 (Bloomberg)

The forecasting power of the Conference Board Leading Economic Index (LEI) has also been researched. Ozyildirim, Schaitkin and Zarnowitz (2010) estimated the validity of the leading index and applied its methodology for euro area. They used the autoregression of euro area CEI as a benchmark to evaluate wether adding the euro area LEI to the estimation would reduce out-of-sample forecasting errors. Their key finding was that adding a LEI to the CEI estimation will decrease the amount of forecasting errors and it is statistically significant.

LEI is not the sole leading index with previous research. OECD publishes their own combined leading index called Composite Leading Iindicator (CLI). CLI has been used to forecast US industrial production. Francis Diebold and Rudebusch (1991) were not able to show any statistically relevant results in their study for predicting Industrial production with CLI. In fact, the univariate autoregression of the IP itself showed more reliable results than the forecasted estimation including CLI. However, there are also more recent studies on where CLI is used to forecast US industrial production. Heij, Dijk and Groenen (2009) used CLI to forecast IP and they were able to show positive results where adding CLI redusesd the amount mean squared errors 12% on a monthy basis compared to the alternative autoregressive approach.

3 BALTIC DRY INDEX

Baltic Dry Index is a maritime freight index for dry raw materials. It has been calculated daily since 1985 and the index compounds weighted average cargo prices for major global shipping routes. The index is owned and published daily by the London Baltic Exchange who receives the shipping quotes from the shipbrokers (Baltic Exchange, 2019). In addition, the major financial news outlets often consider the Baltic Dry index as a leading indicator of stock market. (Wall Street Journal, 2019).

The market share of dry bulk has been reported to contain over 70% of the total world's total amount of freight cargo (Review of Maritime Transport, 2014) and Baltic Dry Index can then be argued to represent major shipping fluctuations. Low level of the index indicates that there is an imbalance for the freight, and it is usually assumed that the weak levels of economic activity cause this imbalance. High levels of BDI would work the opposite way and indicate that there is an increase in the economic activity.

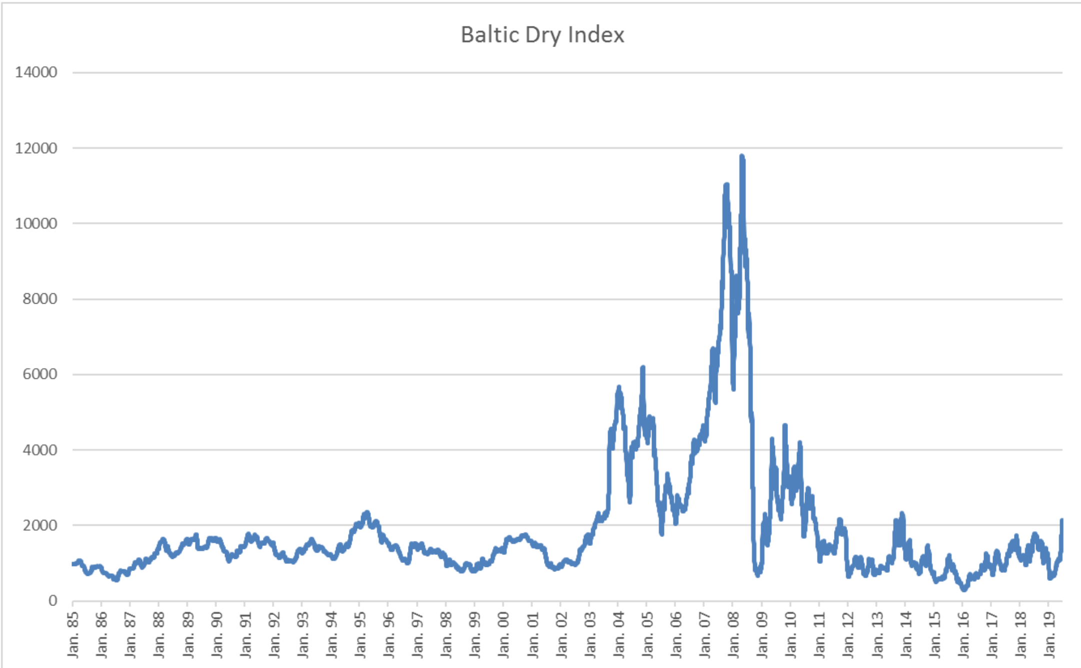


Figure 5 Baltic Dry Index 1985-2019 (Bloomberg)

3.1 Decomposition of BDI

The three shipping classes included in the Baltic Dry Index are Capesize, Panamax and Supramax. Capesize refers to a charter that is too large to access neither canals of Suez nor Panama and therefore requires to use longer freight routes. Cape Size freight class is above 150 000 dead weight tons (DWT). Panamax ships can carry over 65 000

DWT of cargo and respectfully refers to ships, which can still access the Panama Canal. Supramax is a smaller carrier class and has a capacity of 50 000 – 60 000 DWT. (Maritime Connector).

Baltic Dry Index originally also included smaller (15 000 – 35 000 DWT) Handysize carriers but the current constituents of the index do not include Handysize class anymore. Some commodity exchanges such as Chicago Merchandise Exchange – CME (CME Group, 2019) and European Energy Exchange – EEX (European Energy Exchange, 2019) also list derivatives based on Baltic Dry Index constituents. Derivatives have been created to facilitate risk diversification among the market participants.

3.1.1 Index methodology

The Baltic Dry Index consists of 40 percent Capesize vessels. 30 percent Panamax size vessels and 30 percent of Supramax vessels. It compounds together the weighted average of the time charter prices daily. The brokerage members of the London Baltic Exchange will get the prices from their agents looking for shipments and the price of the index is quoted 13:00 GMT. Baltic Dry Index consists prices limited to 22 major shipping routes. (Baltic Exchange, 2019)

3.2 Research of the index

There has been previous research of the Baltic Dry Index and especially its relation to commodity prices. It could be logical to think that shipping prices would drive commodity prices but Gu, Zhenxi and Lien (2018) were able to shed some light to iron ores price effect on Baltic dry index. More specifically they studied the price relation between the iron ore spot price and Baltic dry index levels. Gu, Zhenxi and Lien (2018) concluded that it is the demand rather than supply what drives the iron ore market. Especially the Chinese demand of iron ore in the commodity affected significantly to the dry bulk shipping price of which the Baltic Dry Index represents. Around 29 percent of overall dry bulk shipped is Iron ore (Review of Maritime Transport, 2017) and therefore the study conducted by Gu, Zhenxi and Lien (2018) is relevant when price relations are considered.

Also, various efforts have been made to use Baltic Dry Index as a predictor of global stock returns. Bakshi, Panayotov and Skoulakis (2011) were able to construct an econometric study and show that an increase in the Baltic Dry Index could indeed lead to a statistically significant increase in the global stock markets. Bakshi, Panayotov and Skoulakis (2011) argue that the predictability of the stock returns stems from the developments and implied expectations in the real economy and therefore would satisfy the efficient market hypothesis (Fama H. E., 1965) as the share prices would then represent the fundamental value of the asset price itself. However, Alizadeh and Muradoglu (2014) concluded in their study that the forecasting power for US Stock returns was positive only when there was no significant excess demand nor supply in the global shipping market. As represented in the Figure 5, the equilibrium of shipping prices was clearly not met during the time preseding and succeeding the financial crises from 2007 to 2009. Result of loss in predictability of stock returns during imbalance in equilibrium was also found in smaller raw commodity producing markets. Baltic Dry Index showed significant predictative power for South African stock returns from 1985 to 2008, but lost its predictative nature during the shipping glut of 2008. (Sartorius, Sartorius, & Zuccollo, 2018).

However, the researchers have also found evidense of long term predictability of Baltic Dry Index over time for financial asset prices. Nicholas Apergis and James E. Payne (2013) showed that Baltic Dry index had significant predictive power developed market stock returns. Not only it had predictable power for the share prices but it also performed better when compared to oil prices and MSCI world index. They highlighted that Baltic Dry index was especially better predictor of asset prices to other non-traditional leading index and it made the predicitive power of MSCI world and oil prices significantly stronger over a longer period.

3.3 Advantages of the BDI

One possible advantage of the Baltic Dry Index could be that it is relatively difficult to speculate. Unlike stock market indices such as Dow Jones, Eurostoxx 50 or S&P 500, Baltic Dry Index levels are determined by the physical shipping prices. Due the slow building time and capital-intensive nature of the vessels, supply of ships is inelastic in the short term. Therefore, it supports the argument of lesser speculative nature of the index compared to the paper traded financial instruments.

In a general discussion it has been repeatedly stated that this physical nature of the constituents would reflect a more realistic view of the current state of the trade than indices based on trades on financial instruments. Geman and Smith (2012) presented an analogy that if the supply is correctly assumed as fixed and if there is more demand for the cargo than there are ships, the shipping prices would go up. The relation works also the other way around and the weaker demand would force the ship owners to reduce prices to compete for the orders. This cargo demand is then reflected in the Baltic Dry Index as total daily weighted average shipping price.

3.4 Limitations of the BDI

Even if the relatively slow ship building time could be argued to clean the index out of speculative tendencies from the supply side it also could lead to weaker price discovery. Papailia, Thomakos, & Liu (2017) noticed also in their research of the cyclical properties of BDI that there are challenges especially with the production lag of new ships which means that the supply of cargo ships is very inelastic. On the other hand Papailia, Thomakos, & Liu (2017) found out that the demand for the vessels is extremely elastic.

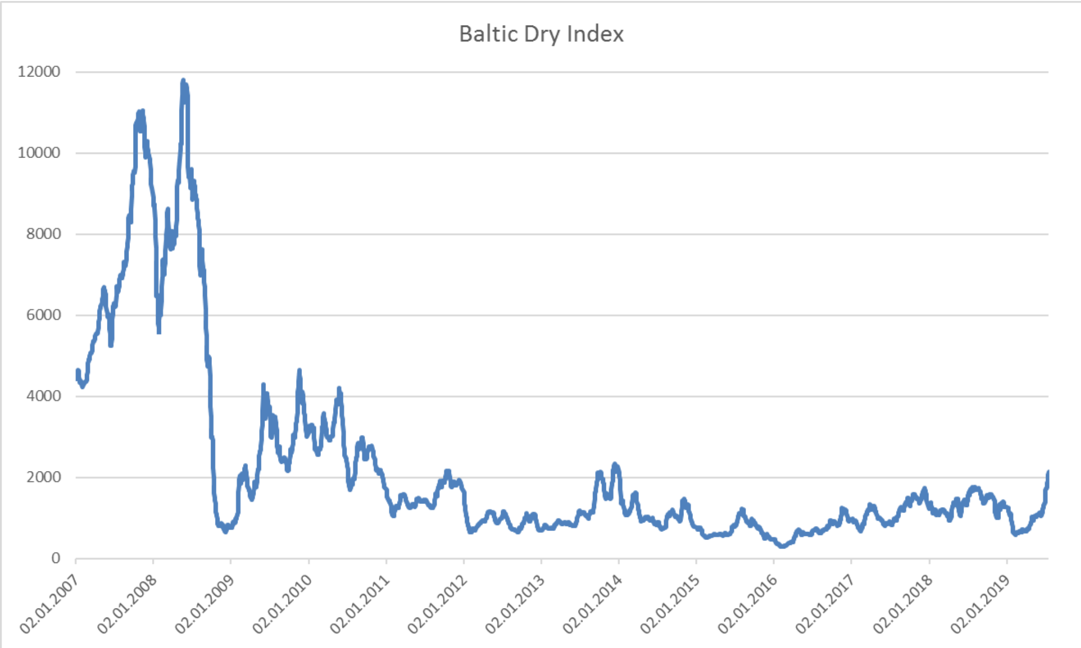


Figure 6 Baltic Dry Index 2007-2019 (Bloomberg)

Traditionally the equilibrium imbalance is thought be generated on the demand side of the shipping industry since ships are relatively slow to be built (Papailia, Thomakos, & Liu, 2017). However, this does seem to be the case in past years. After 2009 Dry Bulk prices have fallen significantly due global shipping glut released to the market. According to industry reports (United Nations, 2009) the glut was developed due an oversupply caused by increased number of vessels. Major reason is that before the 2008 financial crisis the ship building industry experienced extremely high demand, which lead to significant amount of orders. These orders were signed in the overheated world economy, which then took a global down turn due financial crisis and the vessels were effectively finished and delivered at time of weaker global economy (United Nations, 2009). Similar fall of the shipping prices is also seen in BDI (Figure 6). The physical edge against of financial paper indices seemed to work against the BDI itself due the production lag. The turning point of ever lowering shipping prices was not until March 2016 when Baltic Dry Index hit the record low 429 points (Figure 6).

4 DATA AND METHOLOGY

This chapter outlines the methodological framefork for the study. The intention is to find whether Baltic Dry Index (BDI) adds value to forecasting US Industrial Production (IP). The distributions are characterized and stationarity is tested to begin with. Recursive 1 month and 6 months simulation out of samples forecasts of IP are chosen as a method to evaluate the applicability of leading indidicators itself and if Baltic Dry Index could add the forecasting power of the currently used indicators. US Industrial Production (IP) is estimated with the following time series variables being used:

- Conference Board Leading Index (LEI)
- Baltic Dry Index (BDI)
- US 10y3m government bond yield spread (SPREAD)

Conference Board Leading Index already includes the yield spread and therefore series of LEI and SPREAD are not used within the same model. The reason why SPREAD is used in the estimation separately is to gain perspective for the predictive qualities of the data used within the leading index.

Once evaluating the statistical properties in a form of distribution overview and stationarity, general models of the estimations and forecasts will be introduced. Regressions of IP are estimated with lags of 1 and 6 months with expanding estimation window from January 2010 to November 2018. Data sample for the estimations is monthly observations from January 1985 to December 2018. Forecasts are built based on the estimations recursively 1 and 6 months ahead respectively. The aim is to compare forecasted and actual values of IP from January 2010 to December 2018. expanding estimation window with one-step-ahead forecast is endorsed to ensure the robustness of the empirical results.

To gain more information, recursive multivariate forecasts are compared against to the recursive univariate forecast of IP. Univariate forecast model is created for the purpose of asserting information of LEIs forecasting power in general. LEI is applied in a similar manner to estimate and forecast IP as Christiaan Heij, Dijk van Dick and Patrick Groenen conducted OECDs Composite Leading Indicators (CLI) in their study (2009).

In order to evaluate if momentum enhances the accuracy of LEI, univariate and multivariate models are conducted simultaneously for IP in a form of autoregressive with exogenous input model (ARX). The estimation methodology follows the same expanding estimation window approach as the before mentioned models. One-step-ahead forecast models are constructed in similar recursive manner. Evaluation criteria for in sample forecasting is detailed after the introduction of the estimation and forecast models. ARX-model follow the approach of

4.1 Statistical properties

Properties of the time series are inspected before empirical use. Distributional properties affect to the outcomes and the reliability of the estimation and therefore they are briefly introduced. Moreover, stationarity of each time series s observed to ensure the models are using data that can be estimated with the chosen estimation methods. Concept of the basic criteria of stationarity is introduced along with methods of evaluating it.

4.1.1 Testing Stationarity

To estimate a regression, it is generally desirable to have stationary variables. In order to define if the variables used in the model are stationary, I will start with a unit root test. If the unit root hypothesis cannot be rejected the variables are said to be non-stationary. Non-stationary variable then has a unit root and follows a non-predictable Random Walk. Strict stationarity is defined as joint distributions being invariant over time meaning that random variables Y(t + 1)…Y(t + n) have equal joint distribution as Y(t+1+c) …Y(t + n + c), where c is integer. As a condition for economic time series strict form of stationarity is quite restrictive and therefore it is commonly accepted to use weaker form of stationarity as a requirement for estimation. (Tsay, 2010)

E(Yt) = μ = constant, ∀ t
Var(Yt) = σt2 = σ2 = constant, ∀ t
Cov(Yt, Yt-j) = σtj = σj = constant, ∀ t

Purpose of this study is to discover reliable econometric results and therefore it is desirable to be able to reject the null hypothesis stating that the variables used in the regression are non-stationary. As it is often done in financial econometrics, I accept weak form of stationarity meaning that the statistical properties of the series such as mean, variance and covariance are constant over time (Tsay, 2020). In this thesis, Augmented Dickey Fuller unit root test is applied, because of using economic variables that would most likely have serial correlation among each other. (Dickey & Fuller, 1979). If the series used would not fulfil the criteria of weak stationarity the results of the estimations are unreliable. Therefore, it is adamant to avoid non-stationarity in the model because it can lead to spurious regression.

4.1.2 Distribution

In order to execute an adequate econometric study, one must take a closer look at the variables used in the estimation. Mean, Variance, Skew, and Kurtosis of each temporal series are calculated to provide an overview of data being used.

Even if the estimation methodology does not necessary require a normal distribution within the data, it would be the preferred to have a normally distributed dataset. However, possible limitations that the abnormal distribution could cause are accepted.

4.2 Linear Estimation

Ordinary least squares method is used to estimate fluctuation of the United States Industrial Production (IP) from January 2010 to December 2018 on a monthly basis. The intention is to compare the predictive power of Conference Boards leading index (LEI) and the added component of Baltic Dry Index (BDI). Additionally, IP is estimated with the US government yield spread (SPREAD). IP is estimated with using lags of 1 and 6 months in separate estimations. The complete sample for the data is from January 1985 to December 2018.

Also, a univariate model is constructed as a comparison to gain perspective of the usefulness of the leading indicators in general. This is due to a conflicting preceding research of Diebold and Rudebusch (1991) where it was discovered that the univariate autoregression model of IP had more forecasting power than the respective OECDs Composite Leading Indicator CLI. The univariate model is used as a reality check to assess the meaningfulness of using the leading indicators and also as a comparison benchmark.

Both multivariate and univariate models are estimated repeatedly over the estimation window from January 2010 to December 2018. The expanding expanding window approach is used in the regressions meaning that each regression includes the latest available data point available at the time.

4.2.1 Univariate models

As stated earlier in this chapter, the models are estimated using expanding window from January 2010 to November 2018. Univariate model of IP uses a dataset from January 1985 to December 2018. To decide an optimal univariate construction, it is adamant to pay attention to the autoregressive (AR) processes. The construction is chosen at the end of the data sample in December 2018. After inspecting the individual AR processes, based on the partial correlation of IP, there is enough information to choose the right lag construction for Autoregressive model. Additionally, Akaike information criterion is evaluated with the respective lags. AR-model in an expanding estimation window scheme t= 1, …, T (Hyndman & Athanasopoulos, 2018) is presented below:

$$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_n X_{t-n} + \varepsilon_t$$

Parameter φ denotes the autoregressive terms with the with the lag term *n*. The regression is based purely on previous values of the dependent variable *X*.

4.2.2 Multivariate models

The general form of the estimation process with estimation window scheme t= 1, …, T is written below, and each estimation of IP will follow the same expanding approach. The *h* notation refers to the lag term and will correspond with the forecasting horizon later on.

$$Y_t = \beta_0 + \beta_1 x_{t-h} + \varepsilon_t$$

The objective is to measure the effect of having the Baltic dry index as a component in the forecast. At first, I will run a series of regressions from January 2010 to November 2018 where only LEI is used as explanatory variable for the IP. The estimation is performed with lags of 1 and 6 months separately. Lag period of the variable is indicated with the notation *t-1* for 1-month and *t-6* for the 6-month lag. Estimation equations at a single point of time can be expressed via the two following equations below:

$$IP_t = \beta_0 + \beta_1 LEI_{t-1} \tag{1}$$
$$IP_t = \beta_0 + \beta_1 LEI_{t-6} \tag{2}$$

In the second phase, Conference Board Leading index (LEI) and the Baltic Dry Index (BDI) are used as explanatory variables to estimate the Industrial Production (IP). Again, both are estimated with lags of 1 and 6 months recursively over the time frame from January 2010 to November 2018 separately as detailed in below estimation equations at a single point of time.

$$IP_t = \beta_0 + \beta_1 LEI_{t-1} + \beta_1 BDI_{t-1} \tag{3}$$
$$IP_t = \beta_0 + \beta_1 LEI_{t-6} + \beta_1 BDI_{t-6} \tag{4}$$

Additionally, I am estimating IP with the yield spread notated as SPREAD. The aim is to run separate estimations to compare the results of adding the BDI component to the regression. The regression can also be written as followed. As it is with the previous equations, lags of 1 and 6 months are used over the expanding estimation window from January 2010 to November 2018. Estimation equations of the US government bond yield spread at an individual juncture of time are as followed:

$$IP_t = \beta_0 + \beta_1 SPREAD_{t-1} \tag{5}$$
$$IP_t = \beta_0 + \beta_1 SPREAD_{t-6} \tag{6}$$

All above multivariate estimations of IP follow the same expanding estimation window where the latest available data is used to estimate the variable. This will lead to in total of 6 different estimation equations and each equation is estimated 107 times over the estimation window. The expanding estimation windows are used to build a recursive forecast for each model.

4.2.3 ARX-models

Univariate and multivariate approach can also be combined. To see if momentum has positive impact to the estimation an autoregressive with exogenous input, shortened as ARX, model is chosen. If autocorrelation is in place, IP can be estimated linearly with AR-process and external variable. ARX-model follows the approach presented by Stock & Watson in their 2003 publication of leading indicator forecasting. General form of ARX-estimation model is presented below with expanding estimation window t= 1, …, T.

$$X_t = \beta_0 + \beta_1 x_{t-h} + \varphi_1 X_{t-h} + \varepsilon_t$$

The *h* notation refers to the lag term of the exogenous variable. One month lagged LEI is chosen as exogenous variable and AR-process is chosen based on the univariate model selection. Lag term will correspond with the forecasting horizon of the model.

4.3 Forecast

Simulated out of samples, abbreviated commonly as SOOS, method is used to forecast US industrial production. After estimating regressions of IP with an expanding window from January 2010 to November 2018 with a dataset from January 1985 to December 2018, forecasts for each estimation are generated recursively from January 2010 to December 2018. Estimations with 1-month lag are forecasted one period ahead and estimations with 6-month lags are forecasted 6 months ahead. Forecasted values are then compared to the actual values. The purpose for this is to further evaluate the properties of the different models for forecasting purposes. In other words, I compare which forecast is closer to the actual values to find evidence whether the added component of BDI enhances the forecasting power or not. Forecasts $\widehat{Y}_{t+h}$ are based on the estimated coefficients of $\widehat{\beta}_n x_t$ with estimation window being t = 1, …, T. To conduct a forecast, the error term ε_t is set as null (Hyndman & Athanasopoulos, 2018). The general form of the forecast model is presented below.

$$\widehat{Y}_{t+h} = \hat{\beta}_0 + \hat{\beta}_1 x_t$$

The forecast horizon of the above estimation is notated with the letter *h*. Each forecasted estimation within the estimation window *t* = 1, …, T are evaluated against the actual values. Once each forecast estimation can be gauged against the actual outcome, the performance of each forecast model within the whole estimation window *t* can be evaluated. The interest of the forecast performance lies within the differences between the forecasted models. If the forecast performance of US Industrial production (IP) would be better when additional parameters to Conference Board leading index (LEI) are added, it could bring out evidence that the added parameter could work as a part of the leading index as well.

Univariate estimation of IP is also forecasted one step ahead and compared to the actual values. Equivalent recursive approach is applied and the same forecast estimation window t = 1, …, T is used. General form of the recursive forecast equation is formulated followingly:

$$\widehat{X}_{t+h} = c + \hat{\varphi}_1 X_t + \hat{\varphi}_2 X_{t-1} + \cdots + \hat{\varphi}_n X_{t-i}$$

The one-step-ahead approach is chosen to compare the general performance of LEI as a leading indicator to the autoregressive approach where no outside input is used. In addition to multivariate and univariate forecasts, a combined approach is used in form of ARX forecast. Forecast is based on the ARX-model where univariate AR processes and external inputs are estimated together. General form of the ARX-forecast model is presented below. Recursive one-step-ahead approach is applied and the expanding estimation window t = 1, …, T is used.

$$\hat{X}_{t+h} = \hat{\beta}_0 + \hat{\beta}_1 x_t + \hat{\varphi}_1 X_{t-i}$$

The reason behind applying the combined approach is to further evaluate whether leading indicators provide valuable information within themselves when compare to the purely univariate method.

4.3.1 Evaluation of the forecast

I use root mean square error, mean absolute scaled error and Theil inequality coefficient as a criterion to evaluate the forecasts of the estimations. Low value as the root mean square error, abbreviated as RMSE, indicates that the actual and the predicted value of the forecasted value are close to each other (Barnston, 1992). High value as root mean square indicates larger gap between the forecasted values and therefore low value is desirable (Holmes, 2000). More general form of the RMSE for the forecast is shown underneath. $\hat{Y}_{t+h}$ represents the forecasted value and Y_{t+h} the actual value, n being the sample size.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_{t+h} - \hat{Y}_{t+h})^2}$$

In addition to evaluating the squared errors of the respective forecasts I am also interested in the absolute errors that occur within the forecasted models. For this purpose, I have chosen to use the mean absolute error, shortened as MAE. It is a useful measure for comparing forecasting accuracy in the case of the IP series as I am evaluating differences between forecasts that have the exact same forecasted variable. However, mean squared error orientated criteria penalizes large scale errors more and therefore RMSE and MSE are often favored.

$$MAE = \frac{1}{n}\sum_{i} |e_i|$$

After assessing the error values of each model, I take a closer look at the Theil inequality coefficients. The value of Theil inequality (*U*) coefficient ranges between 0 and 1 and if U=0 there is no gap between the forecasted and the actual variable (Theil, 1966). In other words, U=0 indicates a perfect fit for the forecast and U=1 indicates worst possible predictive power for the model. The general formulation of the Theil inequality coefficient is presented as follows.

$$U = \frac{\left[\sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_{t+h} - \hat{Y}_{t+h})^2}\right]}{\left[\sqrt{\frac{1}{n}\sum_{t=1}^{n} {Y_{t+h}}^2} + \sqrt{\frac{1}{n}\sum_{t=1}^{n} {\hat{Y}_{t+h}}^2}\right]}$$

5 EMPIRICAL STUDY

The basic research question is this study is to find out whether leading indices possess predictive qualities and if the Baltic Dry Index could be applied as an additional predictor. As mentioned earlier in chapter 3, BDI has had statistically significant predictive power regarding to stock returns in developing (Alizadeh & Muradoglu, 2014) and especially in developed markets (Apergis & Payne, 2013). The aim in this study is to learn if the forecasting power of BDI could be harnessed for real economy as it has been done for the financial markets. Figure 7 offers a basis for the research question as it presents the changes in BDI and US Industrial Production (IP) series in graphical form. At first glance it seems that during significant economic downturns the BDI seems to start the recovery a notch earlier than the real economy. This is shown as an uptick in late 2008 and early 2020. However, an eye-test as a basis of economic research is not sustainable, hence I will proceed with more in-dept analysis.

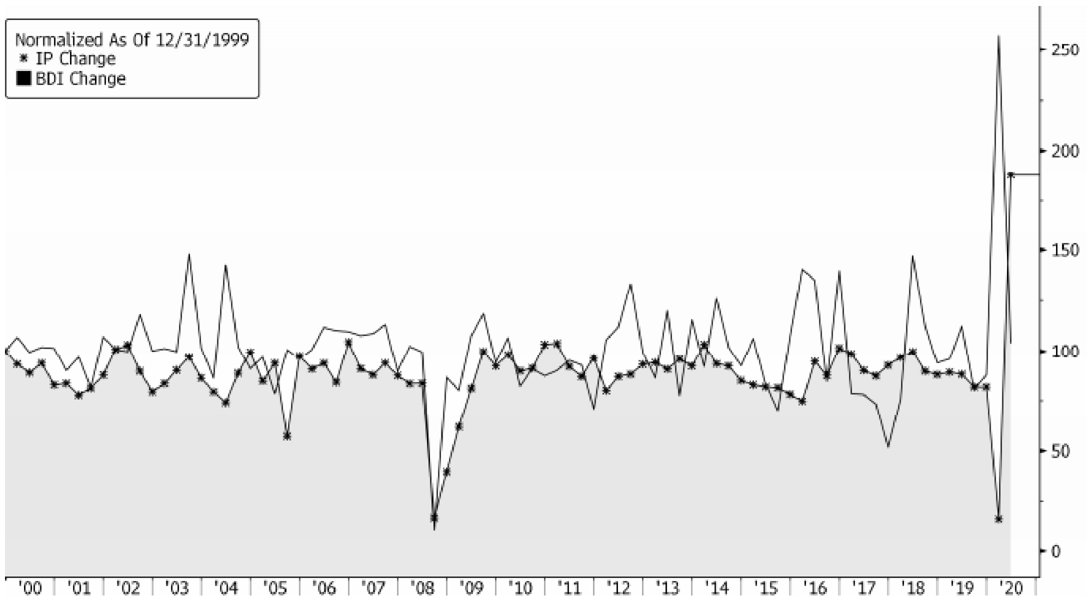


Figure 7 Quarterly comparison of percental changes between Baltic Dry Index (BDI) and US Industrial production (IP) in a normalized scale (Bloomberg)

In addition to analyzing the properties of the BDI, other leading indicators such as Conference Boards Leading Index and the spread in US government bond 10-year and 3-month yields are used evaluate the usefulness of BDI. As mentioned in the literature review: LEI has a multitude of leading indicators combined to form a single leading economic indicator (The Conference Board, 2020) and yield spread has consistently preceded economic slowdowns. One-step-ahead expanding window forecasts are created based on the estimated models and the performance each model within the estimation window is compared against each other. Univariate forecasting model will serve as a minimum criterion and multivariate forecasts of IP should at least outperform to possess predictive qualities. Combined approach is applied when autoregressive with exogenous input (ARX) model is estimated for IP to observe whether momentum could enhance the forecasting power of the leading index.

5.1 Statistical properties

At first, stationarity is tested in this section to evaluate whether the chosen time series are applicable to the estimation or do they need to be altered in a form of differencing or logarithmic scaling. After testing the stationarity of the variables, distributional properties of each variable are characterized and commented briefly. The sub chapter will follow the methodological framework introduced in chapter 4.

5.1.1 Augmented Dickey Fuller Test

As mentioned in the methodology earlier I use Augmented Dickey Fuller test to evaluate the stationarity. Figures 9 and 11 represent the time series used in this study in a graphical format.

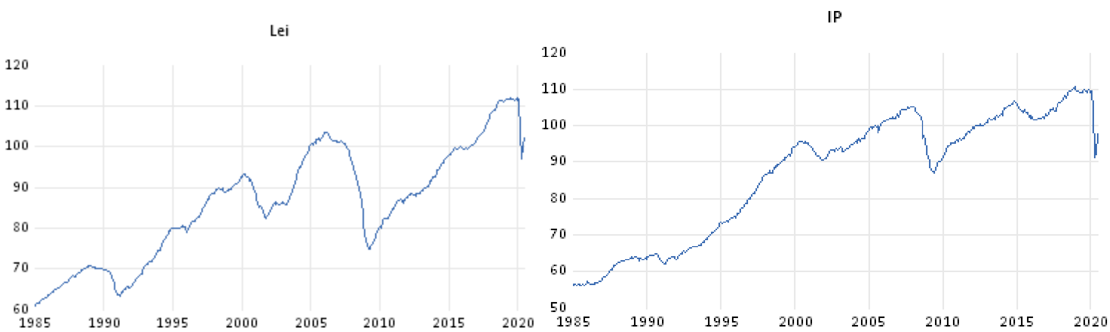


Figure 8 Time series of Leading Index (LEI) and US Industrial Production index (IP) from 1985 to 2020

As shown in Figure 7, the Leading Index and Industrial Production experienced similar incline and decline patterns but the Leading Index (LEI) seems more volatile due its steeper upwards and downwards movements. E.g. It seems that the heavy drops in LEI after the IT bubble of early 2000s and 2008 financial crisis are less sensitive in IP. One argument to support this could be that the financial components in LEI are overemphasizing she severity of the negative shock, which does not snow in IP as it merely indexes the industrial production levels. However, the series themselves are clearly non-stationary due to the fact of a trend being present. Unit root test also supports this, and time series are insignificant with 5% t-test level being -2,009 (LEI) and -1,846 (IP). After differencing and transforming the series to the logarithmic scale, the t-test values of unit root test are -5,019 (LEI) and -5,506 (IP). Both t-test values clearly fulfill the criteria of 5% significance. Differenced graphs of the time series of LEI and IP are shown in figure 9.

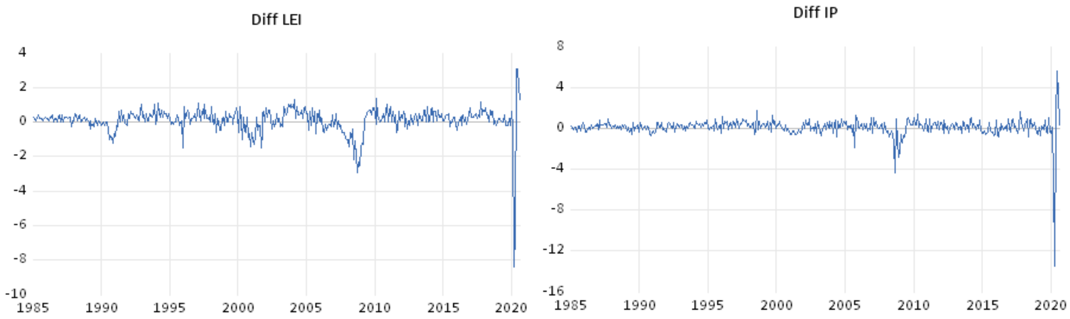


Figure 9 Differenced series of LEI and IP from 1985 to 2020

Baltic Dry Index (BDI) and US Yield Spread are shown in the Figure 10. Unlike with LEI and IP no clear trend is visible. The nature of the measured data explains this behavior and it is expected; BDI measures the shipping price at a certain time and Spread represents the difference of yields in 10-year T-bond and 3-month T-bill. Neither of the values typically accumulate over time and therefore a trend is not expected. However, this does not strictly mean that the series would fulfill the criteria of a weak stationarity. Both graphs in Figure 10 show high volatility which could indicate that the variance over time is not necessarily stable.

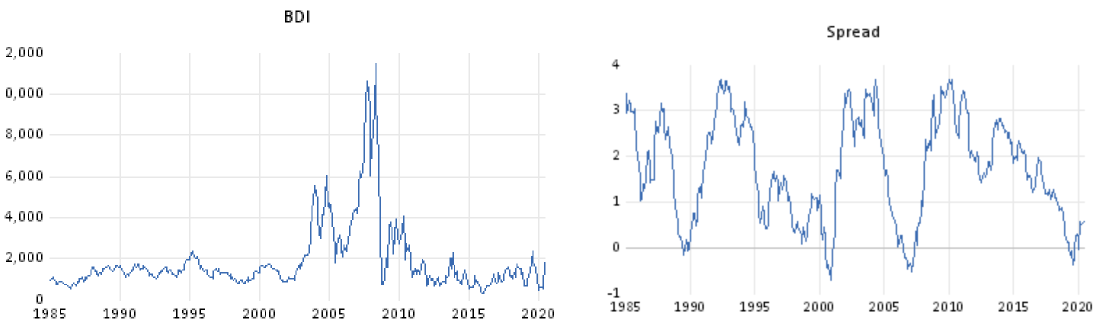


Figure 10 Time series of Baltic Dry Index (BDI) and US Yield Spread (Spread) from 1985 to 2020

After running the unit root test and applying 5% significance the t-test values are -2,929 (BDI) and -3,221(Spread). Both values therefore fulfill the criteria of weak stationarity. However, to smooth out the outliers in the data the differenced series of BDI are chosen to be used. Differenced series are shown in Figure 11.

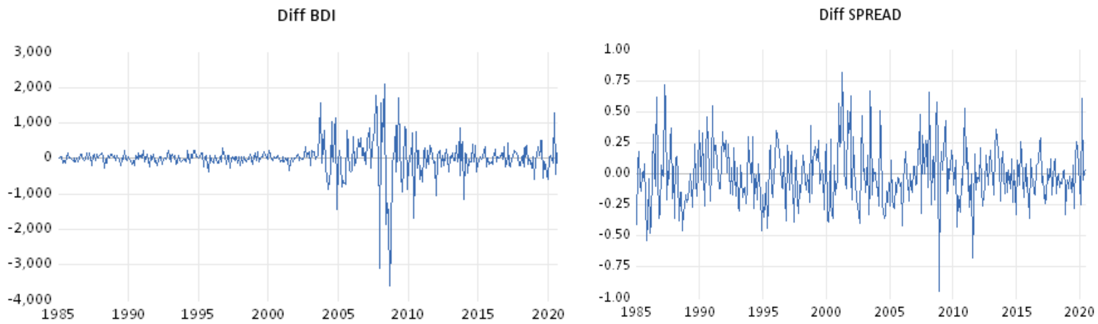


Figure 11 Differenced series of BDI and SPREAD from 1985 to 2020

5.1.2 Distribution of variables

I have composed the distributional properties of the variables used in the estimation in the Table 1. Series of IP and LEI are differenced to satisfy the weak stationarity conditions. BDI and Spread satisfy the conditions of weak stationarity and therefore the temporal series are applicable as they are with 5% significance.

Table 1 Distributional properties of the timeseries

	Mean	Median	Variance	Skewness	Kurtosis
IP*	0.149	0.188	0.353	-0.860	16.775
LEI*	0.081	0.1	0.301	-1.455	7.742
BDI*	0.735	10	221970	-1,797	19.131
SPREAD*	-0.034	-0.01	-0.043	0.362	5.058

**Differenced to the first*

All other variables except SPREAD have a negatively skewed distribution meaning that the median value of the series is greater than the mean. BDI however is skewed to the right and its mean is greater than the median value, hence the positive skew. Kurtosis is imperative to acknowledge since IP, LEI, BDI and SPREAD are clearly leptokurtic with significant excess kurtosis in their respective distributions. Excess kurtosis is commonly defined as kurtosis exceeding the value 3.

5.2 Linear models

As mentioned earlier in chapter 4, a least squares estimation is conducted to measure if Baltic Dry Index (BDI) offers additional explanatory power to estimate US industrial production (IP). The linear estimation is initiated with a univariate model and continued with the multivariate regression on IP. Furthermore, a combined approach is used where autoregressive and exogeneous variables are used to estimate IP. The results of the univariate, multivariate and combined linear regressions are presented in the tables 2, 3 and 4 respectively. The sample size is 408 and it contains monthly values of each temporal series from January 1985 to December 2018.

5.2.1 AR-models

To start univariate estimation, it is essential to analyze the autoregressive (AR) processes of IP at the end of the estimation period in December 2018. As mentioned in the Chapter 4 both autocorrelation and partial autocorrelations of IP are observed to sort out which processes AR-processes are to be inspected with further detail. The graphical representations of autocorrelation (AC) and partial autocorrelation (PAC) in Figure 12 suggest there is higher autocorrelation in the earlier lags of IP with 5% significance. Especially the third lag hints that AR(3) process has explanatory power since both AC and PAC have experience highest spike with the third lag.

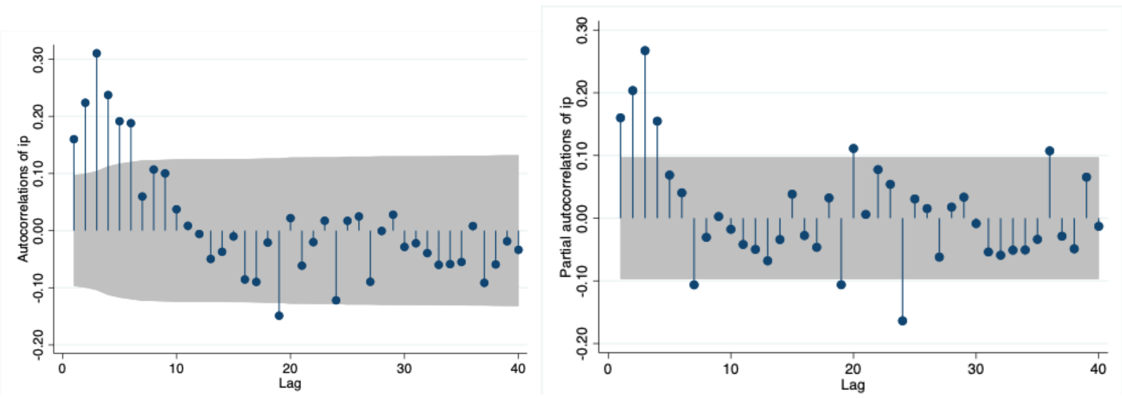


Figure 12 Autocorrelations and partial autocorrelations of IP

Based on AC and PAC representations, the observation of the lags is chosen to be limited for the first 3 AR processes. AR(1), AR(2), AR(3) models are selected for further inspection as well as the model containing multiple AR-processes which are AR(1,2) AR(1,2,3) and AR(2,3). Each autoregressive model is presented in table 2 with the complete sample at the time of last observation in December 2018.

Table 2 Univariate estimation outputs for US industrial Production in December 2018

	AR(1)	AR(2)	AR(3)	AR(1,2)	AR(1,2,3)	AR(2,3)
C	0.1491	0.1492	0.1492	0.1493	0.1494	0.1494
(t-1)	0.1596			0.1271	0.0728	
(t-2)		0.223		0.2026	0.1686	0.1781
(t-3)			0.3087		0.2654	0.2802
sigma2	0.5855	0.5781	0.5638	0.5733	0.5525	0.5540
Log-l	-359.715	-354.880	-344.481	-351.156	-336.203	-337.365
AIC	725.430	715.133	694.960	710.313	682.406	682.730
BIC	737.456	727.160	706.990	726.348	702.451	698.765

AR(3) has the highest coefficient of the observed autoregressive models The model has also the lowest information Akaike information criteria (AIC) scores from the autoregressive models containing only a single autoregressive lag. Additionally, the Bayesian information criteria (BIC), also known as Schwarz criteria, is the lowest for AR(3) when considering models containing a single lag. The estimation output of AR(3) is presented as:

$$IP_t = 0.1492 + 0.3087IP_{t-3}$$

When evaluating AR-models containing more than one autoregressive component the AR(2,3) and AR(1,2,3) seem to have better qualities than AR(1,2). Both AR(1,2,3) and AR(2,3) have almost equivalent AIC and BIC scores but the AR(2,3) is chosen since it has higher coefficient values in the second and third lag. The AR(2,3) model is written below.

$$IP_t = 0.1494 + 0.1781IP_{t-2} + 0.2802IP_{t-3}$$

Both AR(3) and AR(2,3) models are selected to be used as benchmarks in the recursive expanding estimation window forecast. Performance of more advanced forecasting models can be compared to the univariate forecast models as a reality check whether the models containing external parameters are useful at any level.

5.2.2 Multivariate estimation

The table 3 contains estimations for US industrial production denoted as IP_1, $IP_{2,}$ IP_3, IP_4, IP_5 and IP_6 . Coefficients for the intercepts and independent variables are detailed in columns for each model to give an overview of the parameters being used. The estimation output in table 3 is at the time of last available observation within the estimation window.

Table 3 Multivariate estimation output for IP in December 2018

	IP_1	IP_2	IP_3	IP_4	IP_5	IP_6
C	0.1142	0.1227	0.1148	0.1231	0.14816	0.15227
LEI (t-1)	0.4267		0.4190			
LEI (t-6)		0.3684		0.3626		
BDI (t-1)			0.00006			
BDI (t-6)				0.0004		
SPREAD (t-1)					-0.0232	
SPREAD (t-6)						0.0206

The sample for the estimations in table 3 is from January 1985 to December 2018. As the estimation is done on a monthly basis within the expanding estimation window from January 2010 to November 2018, the table 3 results are merely a snapshot of the in single point of time.

LEI with both lags of one and six months has an expected result stating indicating that US industrial production (IP) can be explained by Conference Board Leading Index (LEI). Similar results was found in the previous research when OECDs Composite Leading Indicator was used to as an independent variable to estimate the equivalent dependent variable IP (Heij, Dijk, & Groenen, 2009).

Both BDI and SPREAD have weak effect on IP with the linear regression model being used. With the higher lag term of six month, BDI is even weaker. Spread had a negative correlation with short term lag but with a longer lag term it turns positive. However, with a single point of time estimate there is no basis for a robust analysis of the parameters and therefore the forecast estimations over the expanding window will provide more scientific approach.

5.2.3 Autoregressive with Exogeneous input model

To further evaluate the effect of momentum for estimating US Industrial Production a combined approach between univariate and multivariate estimation is a suitable way to approach the matter. In ARX, abbreviation of autoregression with exogeneous, estimation the dependent variable is estimated in a linear estimation using both autoregressive and external components. Table 4 details the estimation output and the components used to estimate IP with ARX-model.

Table 4 Autoregression with exogenous input (ARX) estimation output for US Industrial Production in December 2018

	$IP_{AR(3)X}$	$IP_{AR(2,3)X}$
C	0.1184	0.1211
LEI (t-1)	0.3732	0.3401
IP (t-2)		0.0903
IP (t-3)	0.1600	0.1801

Two linear estimation models, $IP_{AR(3)X}$ and $IP_{AR(2,3)X}$, are constructed. As shown in the table 4, the difference in the models lies within the autocorrelation parametrisation and both models are using the same 1 month lagged LEI timeseries. AR(3) and AR(2,3) are chosen as autoregressive processes to the estimation due to the selection criteria used earlier. Coefficients and test scores for the processes are presented in table 2.

Ultimately the ARX-estimation is estimated monthly with a expanding estimation window from January 2010 to December 2018. The timeseries used in the estimation sample are from January 1985 to December 2018. A recursive one step ahead forecasting model can be built based on each individual estimation within the estimation window.

5.3 Forecast

After estimating each regression of IP with the expanding estimation window, recursive forecast model is be built. Each regression is forecasted with the respective forecast horizon which corresponds with lag term of the estimation. One month ahead forecast, estimated with 1-month lag terms, is compared to actual values. Equivalently 6-month forecast with lags of 6 periods in the estimation are compared as well to the actual values. In total of 6 different multivariate recursive expanding forecast models are created for the

US industrial production index IP. Furthermore, an autoregressive recursive expanding forecast model of IP is created based on the univariate estimation. Finally, a combined methodology is forecasted with using the ARX-estimation.

Table 5 Multivariate forecast performance of US Industrial Production

	$\widehat{IP}_1^*$	$\widehat{IP}_2^{**}$	$\widehat{IP}_3^*$	$\widehat{IP}_4^{**}$	$\widehat{IP}_5^*$	$\widehat{IP}_6^{**}$
MAE	0.0394	0.0375	0.0367	0.0346	0.0273	0.0537
MSE	0.2698	0.2759	0.2700	0.2778	0.2780	0.2779
RMSE	0.5194	0.5253	0.5196	0.5271	0.5273	0.5272
Theil U	0.6129	0.6272	0.6148	0.6311	0.7475	0.7439

**1 month forecast horizon*
*** 6 month forecast horizon*

Table 5 concludes the multivariate forecast measures. The first finding to pay attention to mean absolute error, shortened as MAE, in first row. $\widehat{IP}_4$ and $\widehat{IP}_5$ have the lowest MAE values, 0.0346 and 0.0273 respectively. However, as the MAE does not punish the accuracy from large scale error it is not used as a sole metric of forecasting accuracy.

In addition to comparing the absolute errors it is also crucial to pay attention to squared errors to reduce emphasize the importance of avoiding large scale forecasting errors. Mean square error, abbreviated as MSE presented in the second row on table 5. $\widehat{IP}_1$, which is the forecast based on IP_1 has the lowest overall MSE value of 0.2698. Consequently, root mean square error, abbreviated as RMSE, value for $\widehat{IP}_1$ is also the lowest.

Moreover, comparing errors between the models it is equally essential to emphasize the importance of the forecasts´ fit. Theil inequality coefficient is presented in the fourth row. Coefficient of 0 would indicate an ideal fit where the forecasted values and actual values match exactly. However, it is rarely the case, but it provides perspective to evaluate the fit as value of 1 would indicate completely unfit forecast. $\widehat{IP}_1$ has a Theil inequality coefficient of 0.6129, which is lower than the equivalent coefficient for any other forecasted multivariate model.

When evaluating the added predictive power of BDI as a component of multivariate regression the results were less distinctive. The explanatory power of BDI is weak for IP and, which reflects to the forecasting results as well. Adding BDI to the forecast model with LEI does not enhance the forecasting accuracy neither with lags of one month nor six months.

Models containing only the yield spread, $\widehat{IP}_5$ and $\widehat{IP}_6$ do not alone as forecast for IP. Their MSE and RMSE values performed close to other multivariate models but the fit is expressively weaker when lags of one and six month are used.

Table 6 includes performance metrics for recursive forecasts for both univariate and ARX-estimations. Expanding estimation window from January 2010 to December 2018 is equivalent as with multivariate forecasts presented earlier. The sample is also same from January 1985 to December 2018.

Table 6 AR and ARX forecast performance of US Industrial Production

	$\widehat{IP}^*_{AR(3)}$	$\widehat{IP}^*_{AR(2,3)}$	$\widehat{IP}^*_{AR(3)X}$	$\widehat{IP}^*_{AR(2,3)X}$
MAE	0.0160	0.0484	0.0662	0.0827
MSE	0.2946	0.3332	0.2588	0.2611
RMSE	0.5427	0.5772	0,5088	0.5110
Theil U	0.6551	0.6336	0.5854	0.5777

**1 month forecast horizon*

Purely autoregressive forecasts $\widehat{IP}_{AR(3)}$ and $\widehat{IP}_{AR(2,3)}$ perform worse than either of the ARX-forecasts $\widehat{IP}_{AR(3)X}$ and $\widehat{IP}_{AR(2,3)X}$ when MSE, RMSE and Theil coefficients are compared. The mean absolute error however is smaller with univariate forecasts, but as noted earlier mean squared errors penalize large scale errors more heavily and are thus favored over absolute errors. Fit of the forecast, measured as Theil U, is also the best when ARX forecasts are compared to the pure univariate forecasts.

When the forecast performance of IP is compared between all forecast models the ARX-forecast stand out as the most accurate forecasts. $\widehat{IP}_{AR(3)X}$, where IP is forecasted with a linear estimation of AR(3) and one month lagged value of LEI, outperforms all other forecasts when mean squared errors and Theil´s inequality coefficient are used as performance criteria. The momentum factor in a form of the autocorrelation parameter seems to enhance the forecasting accuracy.

5.4 Summary of the research results

The empirical results presented in this study support the previous research results of leading indicators stating that leading indicators do have predictive power. The results are aligned with the study conducted by Heij, Dijk, and Groenen (2009) where CLIs predictive power was evaluated for the US industrial production. Unlike in the earlier study made by Diebold and Rudebusch in 1991, the autoregression of the industrial production itself has significantly less predictative power when compared to the multivariate approach where LEI was used as an independend variable to forecast IP. The forecast results of LEI therefore are aligned the study carried by Heij, Dijk, and Groenen (2009) where their CLI forecast outperformed the univariate forecast of IP.

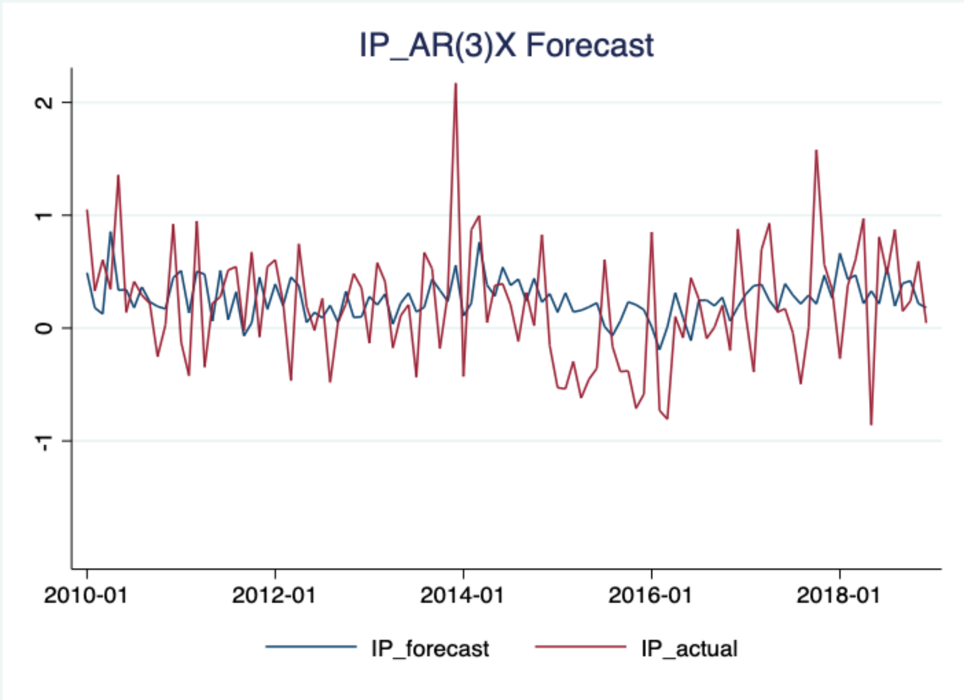


Figure 13 Representation of $\widehat{IP}_{AR(3)X}$ forecast

As mentioned already, in earlier studies leading indices were purely compared to the univariate alternatives. However, the momentum in fact seems to have a positive effect to the forecasting accuracy. The most accurate models used in this study were the ones where autoregressive parameters were used together with exogeneous parameters. Figure 13 visualizes the performance of $\widehat{IP}_{AR(3)X}$ forecast where AR(3) process of IP and one month lagged LEI where used together to forecast performance of IP. This was undoubtedly the most accurate forecasting model used and it suggests that momentum, in for of the autocorrelation, does in fact enhance the forecasting power.

6 CONCLUSIONS

The thesis intended to provide more profound knowledge about relevance of leading indicators. In addition to the traditional leading indicators I was eager to gain more information about the possibility of applying Baltic Dry Index (BDI) as one. The interest behind the BDI lies within its physical nature as it compiles the daily shipping prices for dry bulk cargo. Initially I was determined to understand the leading indicators itself and the theory behind them. Regardless of the popularity in the day to day newscasting, the scientific research of leading indicators appears far from coherent. Outside of the extensive coverage of research related to yield curve, the theory of leading indicators appears opaque as the connection between the gathered index data and the production levels is not very scientific in the traditional macroeconomic sense. The modern macroeconomic theory where the models represent microlevel decision making does not seem to be integral part of the academia and the research is more concentrated on the empirical findings.

Despite the lack of cohesion within the theoretical framework the leading indicators were still able to bring empirical results aligned with recent previous research. I chose to use Conference Boards Leading Index (LEI) because it assembles various individual monthly indicators into a single index figure. The basis of my empirical results was similar as Heij, Dijk and Groenen had in their 2009 study of OECDs Leading Composite Indicator (CLI). I used the simulated out of samples (SOOS) method for the forecast evaluation. US Industrial production was estimated with 10 different models in an expanding estimation window from January 2010 to December 2018. Based on the estimation, recursive one step ahead forecasts were constructed where lag term of the dependent variables corresponds with forecast horizon.

The initial empirical finding was that leading indicators, in this case LEI, do have predictive power over the autoregressive models the production levels. This answers to the first research question whether leading indices bear any significance in the first place. The results are in line with similar study conducted by Heij, Dijk and Groenen in 2009. MSE and fit better for the forecasts using LEI than the autoregressive process. It was explicitly clear with the Conference Boards Leading Index (LEI) I used to forecast US Industrial production, that leading indices do have forecasting power over the univariate approach.

Second empirical finding was that despite the media popularity of Baltic Dry Index (BDI), it does not seem enhance the accuracy of forecasting US industrial production. Models where BDI variable was used alongside with LEI produced less accurate forecasts. This supports the argument that Conference Board Leading Index should not include BDI as a parameter to the index. Baltic Dry Index is not a reliable leading indicator for US industrial production according to the linear estimation conducted.

Regardless of the lag construction used in the estimation, the index did not provide more accurate forecast when compared to the forecasts where only LEI was used. The empirical methods used in this study are rather pragmatic. It could be worth inspecting especially the Baltic Dry Index with more refined approach. Linear methods do not necessarily capture the information in the most ideal manner and non-linear models could be applied. One central limitation is the frequency of the data. BDI has daily values and the major macro data such as the US industrial production is published on monthly basis. This inevitably leads to a situation where all information is not captured. The index however could still be useful for forecasting individual commodity prices as the previous research suggests.

Research covered in the literature review consisted a confrontation between autoregressive models and the leading indicators. Earlier studies conducted either completely rejected (Diebold & Rudebusch, 1991) the accountability of leading indices or endorsed them (Heij, Dijk, & Groenen, 2009). However, the best performing forecast model enabled in the empirical section was $\widehat{IP}_{AR(3)X}$, which consists of AR(3)-process of IP and one month lagged LEI. A robust ARX-model where US industrial production was forecasted recursively over an estimation window from January 2010 to December 2018 had pointedly smaller forecasting errors and better fit than the models where only leading index were used. This opens a new question; should momentum, in a form of autoregression, be included in the Conference Board leading economic index itself as a component? At least, the momentum factor could be a topic for future empirical research on leading indicators.

7 REFERENCES

Adrian, T., & Hyung, S. S. (2008). Liquidity, Monetary Policy, and Financial Cycles. *Current Issues In Economics and Finance*, 1-7.

Alizadeh, A. H., & Muradoglu, G. (2014). Stock market efficiency and international shipping-market. *Journal of International Financial Markets, Institutions & Money*, 445-461.

Apergis, N., & Payne, J. E. (2013). New Evidence on the Information and Predictive Content of the. *International Journal of Financial Studies*, 62-80.

Auerbach, A. J. (1982). *The Index of Leading Indicators: "Measurement without Theory," Thirty-Five Years Later."*. The Review of Economics and Statistics.

Bakshi, G., Panayotov, G., & Skoulakis, G. (2011, January 26). The Baltic Dry Index as a Prector if Global tock Return, Commodity Returns and Global Economic Activity.

Baltic Exchange. (2019, July 19). History.

Barnston, A. G. (1992). Correspondence among the correlation, RMSE and Heidke Forecast. 699-705.

Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 1287-1294.

CME Group. (2019, July). Freight Futures and Options. *https://www.cmegroup.com/trading/energy/freight-futures-and-options.html*.

Cochrane, J. H., & Piazzesi, M. (2005). Bond Risk Premia. *The American Economic Review*, 138-160.

Cook, S. (2019, February). *Forecast Evaluation using Theil's Inequality Coefficients*. Retrieved October 10, 2020, from economicsnetwork.ac.uk: https://www.economicsnetwork.ac.uk/showcase/cook_theil

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

Diebold, F. X., & Rudebusch, G. D. (1991). Forecasting Output With the Composite Leading Index: A Real-Time Analysis. *Journal of the American Statistical Association*, 603-610.

Diebold, F. X., & Rudebusch, G. D. (2013). *Yield Curve Modeling and Forecasting.* Princeton University Press .

Ederington, L. H., & Lee, J. H. (1993). How Markets Process Information: News Releases and Volatility. *The Journal of Finance, 48*(4), 1161-1191.

Estrella, A., & Hardouvelis, G. A. (1991). The Term Structure as a Predictor of Real Economic Activity. *The Journal of Finance*, 555-576.

Estrella, A., & Mishkin, F. S. (1996). The Yield Curve as a Predictor of U.S. Recessions. *Current Issues in Economics and Finance*, 325-339.

European Energy Exchange. (2019, July). Freigh Derivatives. *https://www.eex.com/en/products/global-commodities/freight*.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 383-417.

Fama, H. E. (1965). The Behaviour of Stock Market. *The Journal of Business 38*, 34-105.

Fischer, I. (1930). *The Theory of Interest, as determined by Impatience to Spend Incomeand Opportunity to Invest it.* New York: Macmillan.

Geman, H., & Smith, W. O. (2012). Shipping Markets and Freight Rates: An analysis of the Baltic Dry Index. *THe Journal of Alternative Investments*, 98-105.

Gu, Y., Zhenxi, C., & Donald, L. (2018, December). Baltic Dry Index and Iron Ore Spot Market. *Dynamics and Interactions*, pp. 1-20.

Heij, C., Dijk, D. v., & Groenen, P. J. (2009). *Macroeconomic forecasting with real-time data: an empirical comparison.* Rotterdam: Erasmus University Econometric Institute.

Holmes, S. (2000, 11 28). *Stanford Edu*. Retrieved from https://statweb.stanford.edu/: https://statweb.stanford.edu/~susan/courses/s60/split/node60.html

Hyndman, R. J. (2006). Another Look at Forecast Accuracy Metrics For Intermittend Demand. *Foresight*, 43-47.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice.* OTexts.

Johnson, R. S. (n.d.).

Johnson, R., Zuber, R., & Gandar, J. (2010). *A Re-Examination of the Market Segmentation Theory as a Pedagogical Model* (Vol. 36). Journal of Financial Education.

Kilian, V. &. (2005). *A Practitioner's Guide to Lag Order SelectionFor VAR Impulse Response Analysis.* The Berkeley Electronic Press.

Koehler, A. B., & Murphree, E. S. (1988). A Comparison of the Akaike and Schwarz Criteria for Selecting Model Order. *Journal of The Royal Statistical Society.*(C), 187-195.

Longstaff, F. A. (2004). The Flight-to-Liquidity Premium in U.S. Treasury Bond Prices. *The Journal of Business* , 511-526.

Lutz, K. (2009). Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *American Economic Review*(99), 1053-1069.

Mankiv, N. G., & Summers, L. H. (1984). Do Long-Term Interest Rates Overreact To Short-Term Interest Rates. *NBER Working Paper Series*, 2.

Maritime Connector. (n.d.). Freighter Classification. *http://maritime-connector.com/wiki/ship-sizes/*.

Mishkin, F. S. (1988). What Does the Term Structure Tell Us About Future Inflation? *NBER Working Paper No. 2626*, 2-7.

Mishkin, F. S. (1989). The Information in the Longer Maturity Term Structure about Future Inflation. *NBER Working Paper No. 3126*, 5-8.

Mitchell, W. B., & Burns, A. B. (1938). Statistical indicators of cyclical revivals.

Moore, G. H. (1983). *Forecasting Short-Term Economic Change. Business Cycles, Inflation, and Forecasting* (2 ed.). National Bureau of Economic Research. Retrieved from https://EconPapers.repec.org/RePEc:nbr:nberch:0711.

Nymand-Andersen, P. (2018). Yield curve modelling and a conceptual framework for estinmating yield curves: evedence from the European Cental Bank's yield curves. *ECB Statistics Paper Series*.

Ozyildirim, A., Schaitkin, B., & Zarnowitz, V. (2010). Business Cycles in the Euro Area Defi ned with Coincident Economic Indicators and Predicted with Leading Economic Indicators. *Journal of Forecasting*, 6-28.

Papailia, F., Thomakos, D. D., & Liu, J. (2017). The Baltic Dry Index: cyclicalities, forecasting and hedging strategies. *Empical Economics Springer*, 255-282.

Pindyck, R. S., & Rubinfeld, D. L. (1998). *Econometric Models and Economic Forecasts.* Irwin/McGraw-Hill.

Review of Maritime Transport. (2014). *Developments in International Trade Seaborne Trade.*

Review of Maritime Transport. (2017). *Developments In The Seaborn Trade.*

Sartorius, K., Sartorius, B., & Zuccollo, D. (2018). Does the Baltic Dry Index predict economic activity in South Africa? A review from 1985 to 2016. *South African Journal of Economic and Management Sciences*, 1-9.

Stock, J. H., & Watson, M. W. (2003). *Journal of Economic Literature Vol. 41, No. 3*, 788-829.

Stock, J. H., & Watson, M. W. (2003). How Did Leading Indicator Forecasts Perform During 2001 Recession? *Economic Quarterly, 89/3*, 71-89.

The Conference Board. (2020, May 11). *Calculating the Composite Indexes*. Retrieved from Conference Board Web site: https://www.conference-board.org/data/bci/index.cfm?id=2154

The Conference Board. (2020). *U.S. Business Cycle Indicators.* Retrieved May 1, 2020, from https://www.conference-board.org/pdf_free/press/US%20LEI%20-%20Tech%20Notes%20APRIL%202020.pdf

Theil, H. (1966). Applied Economic Forecasts.

Tsay, R. S. (2010). *Analysis of Financial Time Series.* Chicago: John Wiley & Sons.

Tsay, R. S. (2020, June 15). *Lecture 11: Unit Root and Unit-Root Test.* Retrieved from Chicagobooth.edu: https://faculty.chicagobooth.edu/ruey.tsay/teaching/uts/lec11-08.pdf

United Nations. (2009). *Review of Maritime Transport.* New York: United Nations Conference on Trade and Development.

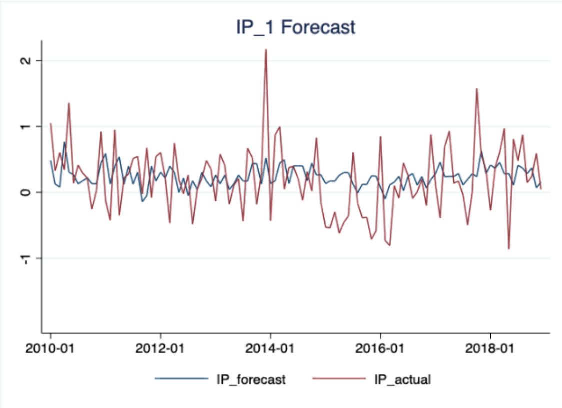
United Nations. (2018). *Review of Maritime Transport.* New York: United Nations Conference on Trade and Development.

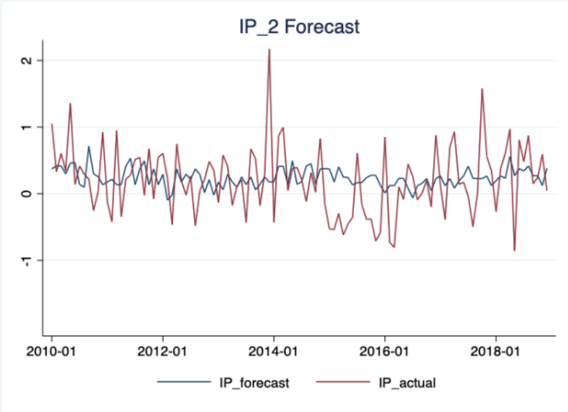
Wall Street Journal. (2019, July 19). Soaring Shipping Gauge Doesn't Spell Growth Boom. *https://www.wsj.com/articles/soaring-shipping-gauge-doesnt-spell-growth-boom-11563537600*.

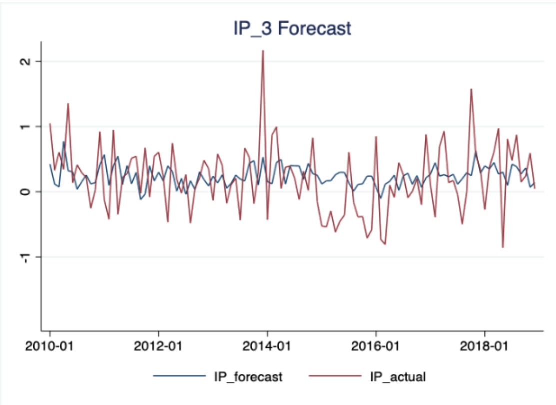
Watson, P. K., & Teelucksingh, S. S. (2002). *A Practical Introduction to Econometric Methods: Classical and Modern.*

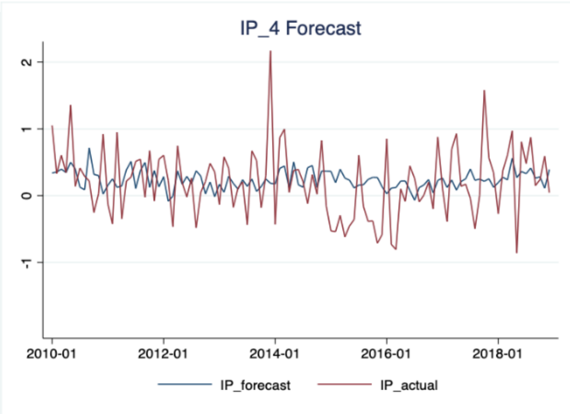
APPENDIX

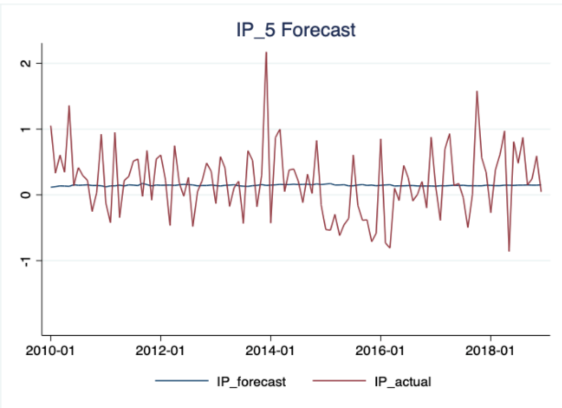
Multivariate forecasts:

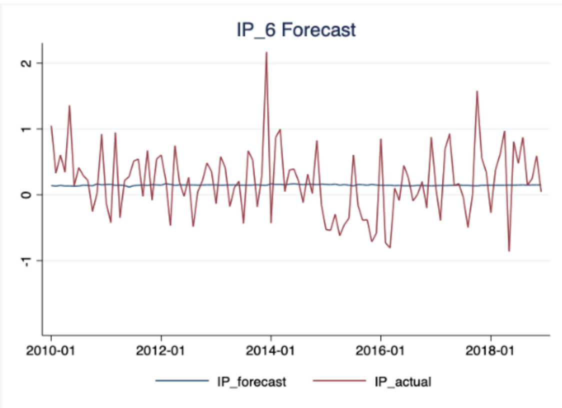




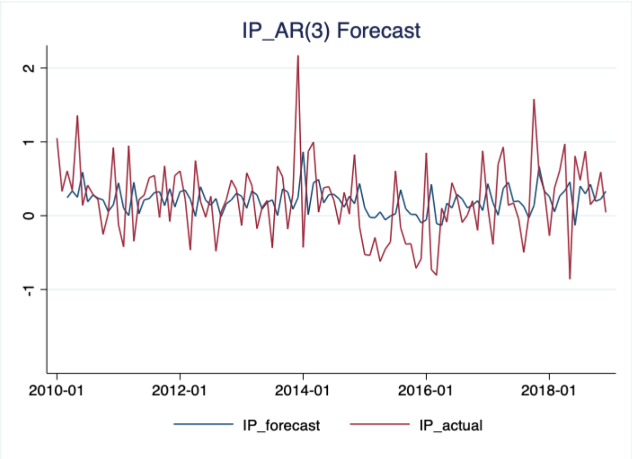


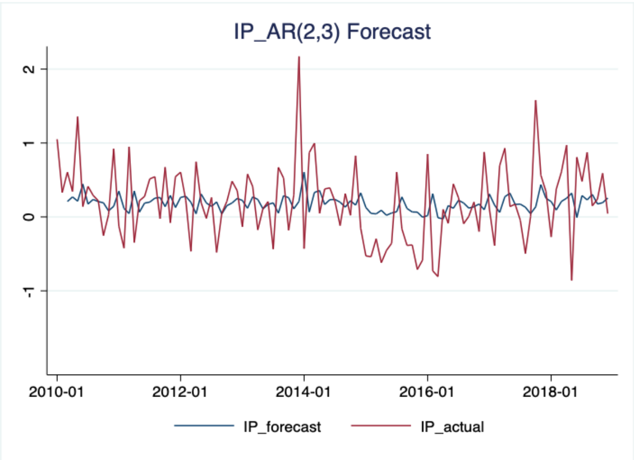






Univariate forecasts:





ARX forecasts: