A Neural Network Approach to Inflation Forecasting

Andrés Adolfo Saenz Guzmán
STUDENT NUMBER: 2049525

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN FINANCIAL ECONOMICS
DEPARTMENT OF ECONOMICS
SCHOOL OF ECONOMICS AND MANAGEMENT
TILBURG UNIVERSITY

Thesis committee:

dr. Gertjan Verhoeven
Prof. dr. Jan Boone

Tilburg University
School of Economics and Management
Department of Economics
Tilburg, The Netherlands
July 2020

To my family,
for their love and support

Acknowledgements

I want to thank my family for their support. They always have been there for me, no matter the circumstances and I know I can count on them.

This thesis was only possible thanks to the constant support of my supervisor, dr. Gertjan Verhoeven. He offered insightful comments and did more than what was expected of him.

Finally, I want to thank all my friends in Mexico and the Netherlands for their help and insights. I want to specially thank Adrian F. Posadas Fuentes, K. Ximena Alfaro Pesina and Ignacio M. García Gallardo for helping me keep my sanity during these times of the COVID-19 pandemic.

Abstract

Inflation forecasting is a fundamental task for central banks and the development of new tools may provide a useful way to attain more accurate predictions. In this paper I compare the performance of a long short term memory neural network with an order one autorregresive model and a vector autoregression for the United States, the Euro Area as a whole, the Netherlands, Chile and Mexico using Python and RStudio[1]*. I find that the neural network outperforms traditional econometric models in three countries, while being similar to econometric methods in one and being outperformed in another, according to mean absolute error metrics.*

***Keywords*:** Inflation, Forecasting, Artificial Neural Networks, LSTM.

[1] I offer the link to the Github repository with reproducible code.

Table of Contents

Introduction

Inflation forecasting is important for all agents in an economy. Investors can make better investment decisions, consumers can improve their consumption smoothing and negotiate better wages and authorities can make better policy decisions. The resurgence of machine learning and neural networks due to advances in technology offer new tools that may be implemented in order to reduce uncertainty by increasing the accuracy of forecasts. There are few studies that focus on inflation forecasting with neural networks, which raises the following research question: Do recurrent neural networks offer an advantage in forecasting inflation relative to econometric autoregressive models?

The structure of the rest of the paper is the following. In section 1 I review the relevant literature, then I explain the methodology and the models used. In the third section, I measure the accuracy of the models and compare them. Finally, in section four are the conclusions and final remarks.

1 Literature Review

To address the literature relevant to this thesis I will divide this section in three categories. First I will explore the literature relative to inflation, since it is the central topic of this paper. In second place, I will review literature on neural networks. Finally, I will approach the literature which combines both topics.

1.1 Inflation

Inflation is one of the main topics addressed by economists and the insights on future inflation are useful not only for policymakers, but for the general public. While Central Banks are interested in producing accurate inflation forecasts in order to design an appropriate monetary policy, individuals benefit from these predictions by allowing them to negotiate better wages and price contracts, to mention a few (Groen et al. 2013). Firms on the other hand, may reduce the uncertainty surrounding the profitability of new projects, since high inflation levels may cause an entrepreneurship unfeasible, this seems particularly true in the case of international joint ventures, where the difference in inflation rates across countries expose firms to additional risks through the exchange rate channel (Kvint 2004).

High inflation is intuitively undesirable. In first place, it has a negative impact in welfare and inequality, given that price rises have a larger effect on middle and deprived classes (Lim and Sek 2015), giving inflation the nature of a regressive tax. Additionally, while it is true that there are no long run effects of inflation on economic growth, in the short run, the neutrality of money does not hold, since high inflation affects real output negatively, which may lead to macroeconomic instability (Faria and Carneiro 2001).

Once established the importance of inflation, it is necessary to explore its determinants as well as the available monetary policy instruments to control it. According to Poole (1970), Central Banks might fight inflation by intervening directly in the money market by adjusting the money supply and buying or selling assets (Quantitative Easing or Tightening) or by changing the market conditions (adjusting the interest rate)[2]. The idea behind these monetary policy instruments are, in the case of quantitative easing (Figure 1), an injection of money supply for the purchase of bonds will increase the price of the safe asset, decreasing its yield, and increase the amount of money in the economy, generating a lower interest rate and higher inflation. This mechanism is based on the quantity theory of money[3] and under the assumptions that in the short run the velocity of money and volume of transactions or output are constant, therefore an increase in the money supply will cause inflation in the same magnitude.

[2] Poole (1970) also suggests that Central Banks may combine both instruments in support of each other to obtain better results. However, this is not always observed.

[3] Fisher's quantity theory of money states that $MV = PY$, where M is the money supply, V is the velocity of money, P refers to the price level and Y stands for output.

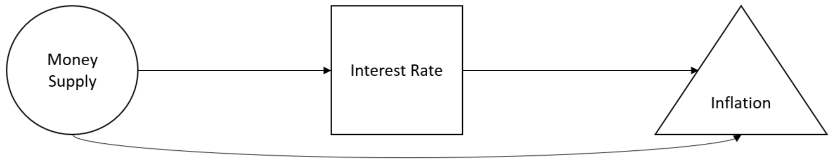

Figure 1: Money supply mechanism. In a circle is the monetary policy instrument, in a rectangle the mediator and in a triangle the outcome variable.

In the case of changes in the money market conditions (Figure 2), an increase in the fixed interest rate would make savings more attractive to consumers than consumption and investment in other assets, thus decreasing aggregate demand and lowering inflation. We can notice that changes in the interest rate also change the cost of present consumption relative to future consumption, making it more expensive to obtain credits and offering consumers cheaper future consumption, lowering aggregate demand and inflation. Finally, a higher domestic interest rate relative to the foreign interest rate would make domestic financial assets more attractive for foreign investors, resulting in a higher demand for these assets and an exchange rate appreciation for the domestic currency, thus decreasing net exports, therefore aggregate demand, and increasing aggregate supply due to relatively lower imported inputs.

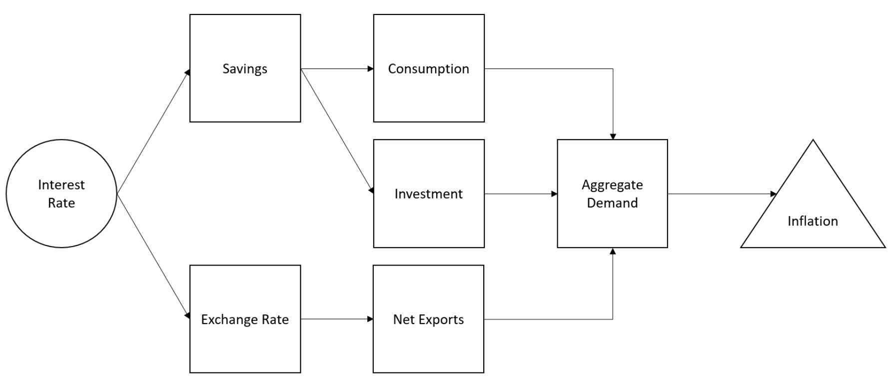

Figure 2: Interest rate mechanism. In a circle is the monetary policy instrument, in a rectangle the mediators and in a triangle the outcome variable.

Economies may present different inflation levels depending not only on inflation targets, but on the characteristics of their markets. The behavior of the goods and services, money, labor and currencies markets may have an impact on an economy's price level. Since the money market has already been covered, given that money supply and the interest rate are the main features of this particular market, I will continue with the goods and services market. Lin and Sek (2015) found that GDP growth has a long run impact in inflation, while the import of goods and services affects low inflation countries negatively due to contagion effects. They also found, regarding the goods and services market, that national expenditure levels also have a long run impact in high inflation countries. Also, it has been noted by Walsh (2010) that increases in real output are usually followed by inflationary periods.

Inflation differentials are also present across countries of the Euro Area, as mentioned by Andersson et al. (2009). These authors show that the position of a country in its business cycle is relevant to explain these differentials. They also suggest that rigidities in product markets may provide a factor to these inflation spreads, since higher regulations in this market raises inflation persistence (Morsy and Jaumotte, 2012). There has also been found that productivity plays an important part on these differentials due Balassa-Samuelson effects (Benigno et al. 2005)[4].

Continuing with labor markets (Figure 3), Andersson et al. (2009) found that wage growth is an inflation determinant, while Morsy and Jaumotte (2012) mention that higher wage rigidities increase inflation persistence. This can be explained in the following manner, larger wage growth and rigidieties may be explained through the presence of stronger labor unions, which give higher market power to workers, whom may overcompensate for expected inflation on their contracts, which leads firms to increase prices in the future, achieving higher inflation persistence. Similarly, unemployment rate plays a main role at the time of establishing wages. A high unemployment rate may translate in lower market power for workers, since firms face low replacement costs (Blanchard et al. 2010), while the opposite also holds true, a low unemployment rate increase workers' market power and replacement costs become higher.

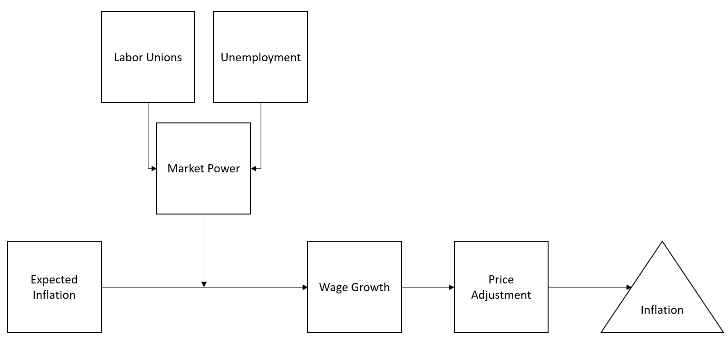

Figure 3: Labor market mechanism. In rectangles are the mediators and moderators, and in a triangle the outcome variable. Note that there is no monetary policy instrument present in this mechanism.

Finally, currencies market may affect inflation due to changes in relative prices of substitute goods. Changes in exchange rates such as a depreciation of the domestic currency may have an inflationary effect on domestic prices (Allsopp et al. 2006). This is due to an increase in prices of foreign goods relative to domestic products, which may lead to a higher demand for national goods overseas, which would increase the domestic aggregate demand and pushing up prices.

It has been noted that differentials in the Eurozone have increased since the European Monetary Union (Pirovano and Van Poeck 2011). These authors hold that these differentials are the result of structural and country-specific factors that, when

[4]The Balassa-Samuelson effect refers to the observation that consumer prices in wealthy countries are higher than those of poorer countries. A usual explanation to this phenomenon is that wealthier countries have a relatively higher non-traded sector's productivity, which lead to a partial price convergence instead of a full convergence.

added to heterogeneous fiscal policies, hinder the ECB's monetary policy, since the monetary authority has a target of 2% inflation in the Euro Area as a whole and does not poses an instrument that allows for multiple targets. However, Morsy and Jaumotte (2012) mention that inflation differentials can sometimes be benign, since they may be a sign of catch up processes, equilibrating mechanisms or temporary shocks.

1.2 Machine Learning and Neural Networks

There are two goals concerning statistical modelling, the extraction of causal information and predicting outcomes given future input variables (Breiman 2001). In order to achieve these objectives two approaches have developed, the first one consist on data modeling, while the other on the creation of algorithms. Breiman (2001) mentions that the data modeling approach consists in the assumption of a stochastic process that may relate input variables with outcomes after allocating some weights to input variables, such that the outcome is a function of predictors, parameters and some random noise. On the other hand, the author explains that the algorithm modeling culture consider the process that relates predictors with the outcome as unknown, thus the objective of this approach is to find a function that is able to relate inputs with outcomes. In order to validate the usefulness of a model, data modeling focus on testing the fit of output and the examination of errors, while algorithm modeling tests the accuracy of predictions. These characteristics mark the main differences between econometric models and machine learning algorithms such as decision trees and neural networks.

Machine learning began to develop during the 1950s due to the surge of an interest in Artificial Intelligence. However, because of the lack of computational power, the advances on the field slowed during the late 1960's until the development of multilayer perceptrons in the 1980s (Boelaert and Ollion 2018). The next moment of high interest in the field came in the 2000s given the improved result of deep learning algorithms when feed with large data sets. The development of artificial intelligence and machine learning has been constrained by the availability of data and computational capability, as well as the technical level of algorithms which have been refined over time. However, the past decade has shown an increase in the interest in machine learning from a more general public, as indicated by the 180,000 data scientist shortage in 2018 estimated by Deloitte in 2016 (Algorithmia 2019).

Although machine learning algorithms originated in Computer Science and Engineering, Economics may benefit from it. As noted by Einav and Levin (2014), the increase in data quality and quantity in recent years offer the possibility to predict aggregated variables such as unemployment, consumer confidence and retail sales, as well as allowing the study of employee and consumer behavior. Research with these large data sets has mainly relied on traditional econometrics, focusing on parameter values and the existence of important omitted variables, while a machine learning approach focuses on predicting future outcomes by selecting the best predictors from the available data. However, machine learning can help improve estimation of causal effects when dealing with high-dimensional data, enhancing the credibility of policy analysis (Einav and Levin 2014, Athey and Imbens 2017). Einav and Levin (2014) hold that over the next decades machine learning tools and big data will play a big role in economic policy and research, complementing common sense and economic theory.

Among the different tools in machine learning, neural networks are one of the most notorious. They consist in mathematical models that seek to resemble the human brain through the connection of artificial "neurons" (Hyndman and Athanasopoulos 2018), allowing them to recognize patterns in data in order to make predictions. The structure of neural networks usually have three components, an input layer, hidden layers and an output layer, although the simplest networks consist only in the input and output layers, resembling linear regressions. Similarly to econometrics, in every layer each input is associated with a weight and a bias[5] is assigned before applying an activation function and sending the output to a hidden layer. At the next stage, the hidden layer will repeat this process in order to obtain more complex relations in the data, increasing the complexity with each hidden layer that is added to the network. Finally, this is sent to the output layer to make predictions in the case of a feedforward neural network[6].

A particular type of neural networks are recurrent neural networks. This subset of neural networks works with sequences of data, meaning that order of inputs matter, let it be text, sound waves or time series. Recurrent neural networks generate a "memory" by incorporating a delay that affects information, allowing the network to generate output that responds to

[5] The bias is known as the constant or interceptor in econometrics.

[6] In the case of back-propagation, information is sent back as input after passing through a loss function in order to produce better results

causality or context from the input signal, the state and the readout[7] of different time steps (Sherstinsky 2020).

There are several applications to neural networks such as text, image, facial and speech recognition, leading to a growing interest to apply these in fields such as medicine and finance (Sirignano and Spiliopoulos 2018). There are different types of neural networks that are better at performing certain tasks. Convolutional neural networks, for instance, are particularly good at dealing with images, text and speech, allowing them to deal with computer vision and natural language problems (Bhandare et al. 2016) such as facial recognition, image classification, action recognition, document analysis and speech and text classification. Other example are multilayer perceptrons, which may be used to predict air quality, since there is a complex relationship between pollution and meteorology (Gardner and Dorling 1998).

1.3 Neural Networks and Inflation

Since the subprime crisis many central banks and regulators have been tasked with supervision and market oversight, giving them access to new data and pushing monetary authorities into big data territory (Chakraborty and Joseph 2017). This has lead central banks to start big data initiatives in order to complement traditional macroeconomic indicators and support central bank policies (Wibisono et al. 2019). The ability of neural networks to capture non-linearities in data give them an advantage at forecasting when variables fail to fulfill the requirement of independent and identically distributed (Šestanović 2019), in other words, neural networks do not require assumptions about the data nor *a priori* knowledge of the functional form that relates inputs to outputs, giving them more flexibility to adapt to training data than traditional econometric models.

The use of neural networks in research to approach inflation analysis has been narrow and the use of recurrent neural networks has been even more limited (Šestanović 2019), even though neural networks seem to perform better at predicting inflation in 45% of 28 OECD countries while AR(1) processes only do so in 21% of them (Choudhary and Haider 2012). Nakamura (2005), Teräsvirta et al. (2005), Binner et al. (2006) and Al-Maqaleh (2016) find that neural networks outperform univariate autorregresive models for inflation forecasting in the United States. However, the author notes that the neural network specification may play a decisive role in its success.

Šestanović (2019) remarks the necessity of identifying the overfitting problem when using neural networks and suggest two possible solutions. On one hand, the problem may be addressed reducing the network's complexity by including only the most important inflation determinants and selecting an appropriate number of neurons in the hidden layers. On the other hand, the use of back-propagation in the neural networks. However, Nakamura (2005) mentions another solution for short run out-of-sample predictions: the use of an early stop procedure, since it prevents the neural network to adjust too much to the training data set. Regarding the structure of networks for inflation forecasting, Šestanović (2019) estimates 250 different Jordan neural networks, a specific type of recurrent neural network, and find that 80% of top ten performers contain two or three hidden layers.

The accuracy of inflation forecasts has a huge impact in monetary policy, making the difference in a central bank's policy decisions. Neural networks offer an additional tool to monetary authorities for times where most econometric models fail due to their linear nature[8], such as periods previous to and during economic crisis (Šestanović 2019). Neural networks and other machine learning algorithms are not going to replace traditional econometrics, but complement them in order to enhance economic research and practice.

2 Models

In this section I present three models for inflation forecasting. The first one is a first-order autoregressive process (AR(1)), which is going to be used as a naive benchmark for other models. The second model is a vector autoregression that takes into account various macroeconomic variables. Finally, I use a recurrent neural network with long short term memory (LSTM) layers to forecast inflation.

[7] The input signal refers to new information, the state to a collection of information created by the layer and the readout to the neuron that produces output.

[8] In this sense, neural networks are more likely to capture outliers when there are non-linear relations in data.

In order to capture the effects that different sectors may have in inflation, I selected variables from the different markets of Hick's IS-LM model, since most macroeconomic models share these markets. From the goods and services market I use the output gap, since it captures the effects of consumption, investment, public expenditure and net exports. To capture the financial market, I use the monetary policy instruments, nominal interest rate and money supply. In order to see the impact of labor market in inflation I use real wages and unemployment rate. Finally, in order to capture possible effects from the currencies market, I selected the real exchange rate (RER).

In order to assess for external validity, I selected five economies with different macroeconomic characteristics. The economies to be studied are the United States, the Netherlands, the Euro Area as a whole[9], Chile and Mexico. I choose these economies because their respective central banks face inflation with different monetary policy instruments. Some favor the quantitative easing approach (Euro Area), while others prefer to fix their interest rate (United States, Chile and Mexico), although both these policies are sometimes combined for a better outcome, following Poole (1970). Similarly, these countries represent developed (United States, the Euro Area and the Netherlands) and developing economies (Chile and Mexico).

Inflation[10] is computed with adjusted Consumers Price Index (CPI) data from each country as usual. However, in order to avoid seasonality problems I use year on year (YoY) inflation instead of month on month. Therefore, I calculate inflation (π) as follows:

$$\pi_t = \frac{CPI_t}{CPI_{t-12}} - 1$$

Since GDP is a quarterly variable, I used each country respective Industrial Activity Index[11] as a proxy for output gap. In order to compute a monthly output gap, and to avoid capturing future information in the time-series I use a one sided Hodrick-Prescott (HP) filter with a smoothing coefficient (λ) of 14,400, as it is the usual value of λ for monthly data in the literature.

As money supply, I use the M2 monetary aggregate from each country's central bank[12], since it captures physical money, traveler's checks, demand and other checkable deposits, positions in money market mutual funds and savings deposits, being a good measure for households held liquid assets. Following Walsh (2010), in order to avoid capturing the effect of "natural" growth of money supply, as depicted in Fisher's quantity theory of money, I use a HP filter with $\lambda = 14,400$, leaving only "surprise" changes in money supply.

For nominal interest rates, I chose the Effective Federal Funds Rate published by the FRED for the United States. For the Euro Area and the Netherlands I use the discount rate for the Euro Area reported by the IMF. In the case of Chile, I chose the monetary policy interest rate announced by Banco Central de Chile. Finally, I used Banco de Mexico's days interbank balance interest rate (TIIE 28) as monetary policy instrument for the country.

For labor market variables, I use the YoY real wage change of each country and seasonally adjusted data on unemployment rate. I use hourly wage rate data from the FRED for the United States and the Netherlands. Given the high levels of informal employment in Mexico, I use average daily wage data from the Mexican Social Security Institution (IMSS), since is the usual for this country. Data for real wage in Chile is build using data from OECD and INE. I use unemployment data for the United States available at FRED's website, while using data from INEGI for Mexico. Finally, I use INE's data published in the Quarterly employment data statistical bulletin (Boletín Estadístico: Empleo Trimestral) for Chile. Unemployment data for the Euro Area as a whole and the netherlands was obtained from the ECB's statistical data

[9] I use the 19 countries (fixed composition) data.

[10] I refer to core inflation.

[11] The data is available for the United States, the Euro Area and the Netherlands is easily found at FRED's website, while the index is can be found at Mexico's Statistics and Geography National Institution (INEGI). Chilean data, however, is harder to find and is available at the Yearbook of National Accounts published by the Statistics National Institution (INE) as IMACEC (Economic Activity Monthly Index) for data previous to 1996 and in its website for data from 1996.

[12] In the case of the Euro Area and the Netherlands I use the same data, since monetary policy should be the same since the monetary union.

warehouse. I was unable to find a long enough time series on real wages for the Euro Area, therefore I will not use this variable in models for this economy.

Finally, I use historical RER data from the FRED for the United States, the Euro Area and the Netherlands. Data for the chilean economy is available at Banco Central de Chile and for Mexico at Banco de Mexico. Similarly to inflation, I use YoY change to avoid seasonality problems in the time-series.

The periods covered in this study go from January 1995 to January 2020 for the United States. From January 1999 to January 2020 for the Euro Area and the Netherlands, because the interest rate data starts in 1999 due to the foundation of the ECB in 1998. The period covered for Chile goes from January 1996 to January 2020, while the range for Mexico starts in January of 2001 until January 2020. Mexican data starts in 2001 because Banco de Mexico's public data on monetary aggregate M2 starts that year.

2.1 A Benchmark

As a naive benchmark, I use a first-order autorregresive model following Nakamura (2005)[13], Binner et al. (2006) and Choudhary and Haider (2012). These processes are often found in empiric literature (Greene 1998), since the use of these models is seen as reasonable when dealing with processes of which there is little knowledge. Additional to this, the prevalence of AR(1) models in the literature are due to its simplicity. The use of more complex models require a correct specification of an usually unknown process, which is in most cases unfeasible.

I specify the AR(1) model as follows:

$$\pi_t = \alpha + \beta L(\pi_t) + \epsilon_t \tag{1}$$

In this equation, π_t is inflation in time period t, α is a constant added due to existence of target inflation rates[14], L is the lag operator, β is the coefficient of past inflation and may be interpreted as past inflation's persistence[15] and ϵ_t is the error term for time period t. Therefore, the regression that has to be estimated is

$$\pi_t = \alpha + \beta \pi_{t-1} \tag{2}$$

I use the 'forecast' package in RStudio by Hyndman and Khandakar (2008) to estimate the model. Since the interest is to produce forecast for a year, I use the last 12 months of each country's data set for testing and the rest to estimate the model as shown in the table below.

Table 1: Split of data for estimating and testing AR(1) model in each country

Country	Estimation	Testing
United States	Jan 1995 - Jan 2019	Feb 2019 - Jan 2020
Euro Area	Jan 1999 - Jan 2019	Feb 2019 - Jan 2020
Netherlands	Jan 1999 - Jan 2019	Feb 2019 - Jan 2020
Chile	Jan 1996 - Jan 2019	Feb 2019 - Jan 2020
Mexico	Jan 2001 - Jan 2019	Feb 2019 - Jan 2020

After estimating the models, I found that inflation processes are stable for each country, i.e. $\beta < 1$. I also note that the value of the constant is consistent with central bank's target inflation range in developed economies[16]. In the United States and the Euro Area target inflation is 2%[17], while the values of the mean ($\frac{\alpha}{1-\beta}$) are 0.0217, 0.0157, 0.0159, respectively. As

[13] Nakamura (2005) uses univariate autoregressive models with 1 to 8 lags.

[14] The AR(1) process converges to $\frac{\alpha}{1-\beta}$, allowing to compare the behaviour of central banks regarding their inflation targets.

[15] Note that in order for the process to be stable, $\beta < 1$ must happen.

[16] Inflation rate targets are usually set at 2% with a 1% error range.

[17] Euro Area's actual inflation target is close to, but below to 2%. However, Hartmann and Smets (2019) mention that the implied target inflation that yields symmetry is 1.76, with a higher and lower bounds of 2% and 1.5%, respectively.

expected, inflation in Latin America is in the higher band of the range because, while target rate in Chile and Mexico is 3%, the values of α are 0.0394 and 0.0482 respectively.

Table 2: AR(1) estimation results and central banks' target inflation for each country

Country	AR(1)	α	Mean	Inflation Target
United States	0.9724	0.0006	0.0217	0.0200
	(0.0131)	(0.0001)	(0.0022)	(0.0100)
Euro Area	0.9660	0.0005	0.0157	0.0176
	(0.0155)	(0.0001)	(0.0040)	(0.0025)
Netherlands	0.9503	0.0008	0.0159	0.0176
	(0.0187)	(0.0002)	(0.0045)	(0.0025)
Chile	0.9763	0.0009	0.0394	0.0300
	(0.0129)	(0.0003)	(0.0108)	(0.0100)
Mexico	0.9707	0.0014	0.0482	0.0300
	(0.0187)	(0.0002)	(0.0072)	(0.0100)

Once estimated the model, I used it to produce a 12 months forecast for the February 2019 - January 2020 period in each country. After computing out-of-sample predictions, I subtracted the real value in order to calculate the model's forecast errors. A straightforward intuition of the error values is positive errors imply that the model overestimates future inflation while negative errors underestimate it.

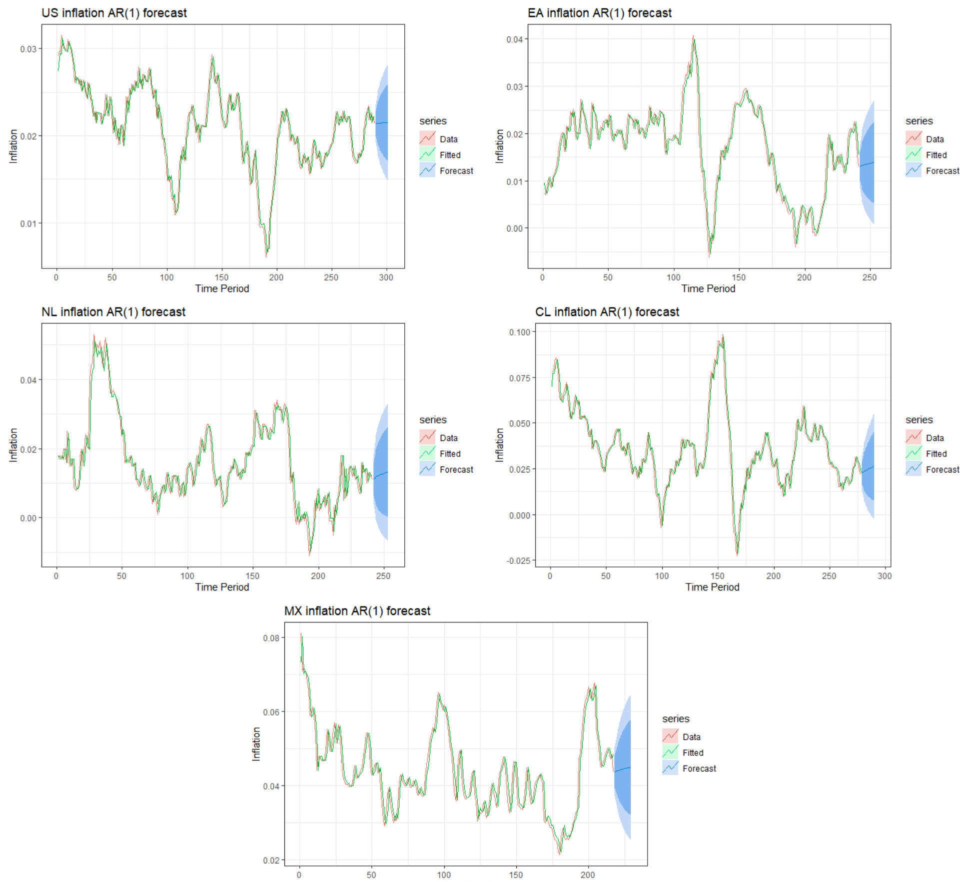

Figure 4: Inflation forecasts from an AR(1) model for different countries. Forecast fans are for confidence intervals of 80% and 95%

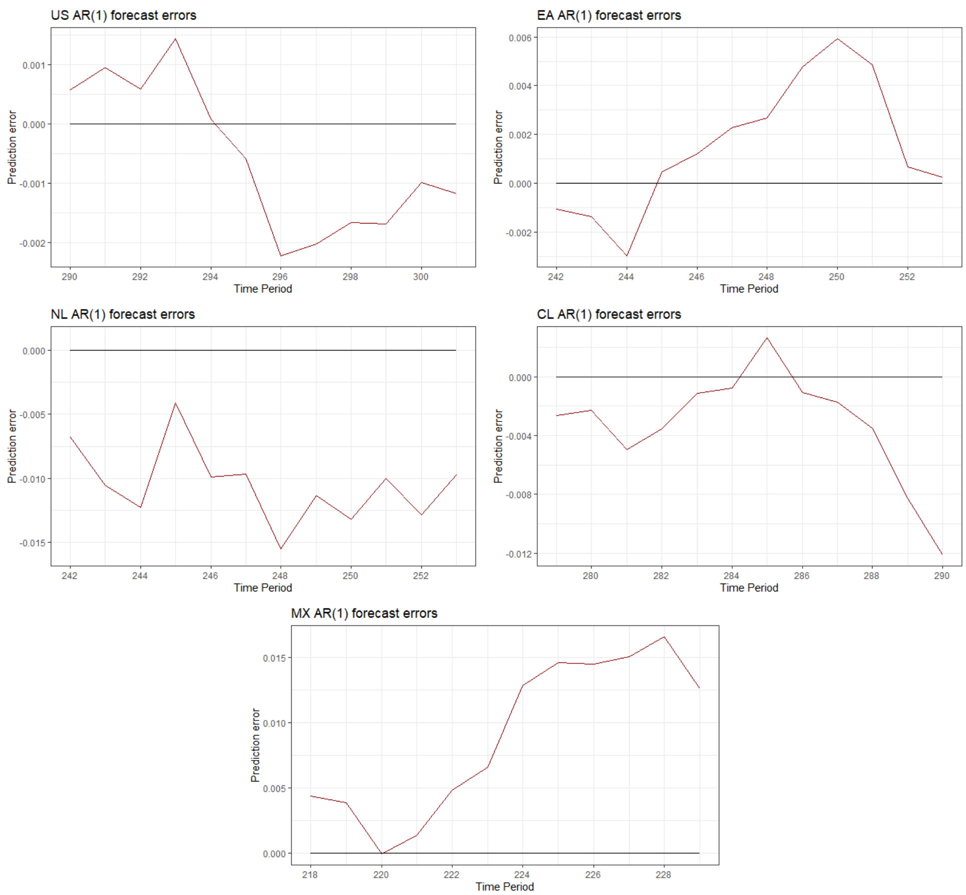

Figure 5: Inflation forecasts errors from an AR(1) model for different countries

As seen above in Figure 5, the benchmark overestimates future inflation for the first five months in the United States, while underestimating it from July 2019 (time period 295) onward. The opposite is true for the Euro Area as a whole. The model underestimates future inflation for the first three months and overestimates the rest. In the case of the Netherlands and Mexico, the AR(1) process underestimates and overestimates future inflation for the whole forecast, respectively. Finally, the benchmark model underestimates Chilean future inflation during the whole forecast except for August 2019 (time period 285).

2.2 Vector Autoregressive Model

In a beginning, central banks used models that captured the effects of several macroeconomic variables in a set of simultaneous equations that were the result of an intersection between macroeconomic theory, mathematics and statistics, these models are known as general equilibrium models. Once defined an appropriate equation system, these equations were estimated, allowing monetary authorities to guide their economies in the desired direction (Cuche-Curti 2006). The success of these models has passed the test of time, although some modifications have been made in order to solve limitations and criticism.

There have been three waves of criticism of general equilibrium models. The first one occurred during the decade of 1950 and 1960 and had to do with their limited empirical methodology (Cuche-Curti 2006). The second wave took place in the 1970s and was about the lack of micro-foundations of the models and is known as Lucas critique. The final wave generated a new type of models for central bankers. The labeling of the division between endogenous and exogenous variables in general equilibrium models as arbitrary paved the path for a new atheoretical model that takes every variable as endogenous, the vector autoregression (VAR). To this day, central banks use different models to study inflation and other variables of interest, being the most popular dynamic stochastic general equilibrium (DSGE) models and vector autoregressions.

The need for central banks for a collection of different models is motivated by the fact that, while VAR models are successful at forecasting inflation, they lack an economic structure, working as a black-box between inputs and outputs and preventing an understanding of the underlying mechanisms. To fight this criticism, these models were endowed with economic theory, giving birth to their structural version (SVAR) and vector error-correction models (VECM). At the end, these models are used along DSGE, the most theoretically advance models available to central bankers, in order to compliment each other and benefit from the best qualities of each type, the study of mechanisms (DSGE, SVAR) and forecasting (VAR, SVAR and VECM).

In this subsection I propose a VAR(p) model using the variables previously discussed, where p is the order of the model. The structure of the model can be described in the following manner:

$$Y_t = \mu + \sum_{j=1}^{p} \Delta_t L^j(X_t) + v_t, \tag{3}$$

where Y_t is a vector of n exogenous variables in time period p, μ a vector of n constant terms, Δ_t an $n \times n$ matrix of parameters, L is the lag operator, X_t an $n \times n$ matrix of variables and v_t a vector of n estimation errors in time period t. Using the proposed variables, the resulting equation system is:

$$\begin{vmatrix} \pi_t \\ i_t \\ M2_t \\ (y-\tilde{y})_t \\ w_t \\ u_t \\ \varepsilon_t \end{vmatrix} = \begin{vmatrix} \mu^{\pi} \\ \mu^{i} \\ \mu^{M2} \\ \mu^{(y-\tilde{y})} \\ \mu^{w} \\ \mu^{u} \\ \mu^{\varepsilon} \end{vmatrix} + \begin{bmatrix} \Delta_{\pi,\pi}^{t-1} & \cdots & \Delta_{\pi,\varepsilon}^{t-1} \\ \vdots & \ddots & \vdots \\ \Delta_{\varepsilon,\pi}^{t-1} & \cdots & \Delta_{\varepsilon,\varepsilon}^{t-1} \end{bmatrix} \begin{vmatrix} \pi_{t-1} \\ i_{t-1} \\ M2_{t-1} \\ (y-\tilde{y})_{t-1} \\ w_{t-1} \\ u_{t-1} \\ \varepsilon_{t-1} \end{vmatrix} + \cdots + \begin{bmatrix} \Delta_{\pi,\pi}^{t-p} & \cdots & \Delta_{\pi,\varepsilon}^{t-p} \\ \vdots & \ddots & \vdots \\ \Delta_{\varepsilon,\pi}^{t-p} & \cdots & \Delta_{\varepsilon,\varepsilon}^{t-p} \end{bmatrix} \begin{vmatrix} \pi_{t-p} \\ i_{t-p} \\ M2_{t-p} \\ (y-\tilde{y})_{t-p} \\ w_{t-p} \\ u_{t-p} \\ \varepsilon_{t-p} \end{vmatrix} + \begin{vmatrix} v_t^{\pi} \\ v_t^{i} \\ v_t^{M2} \\ v_t^{(y-\tilde{y})} \\ v_t^{w} \\ v_t^{u} \\ v_t^{\varepsilon} \end{vmatrix}, \tag{4}$$

where the variables π, i, $M2$, $(y - \tilde{y})$, w, u and ε are inflation, the nominal interest rate, surprise changes in money supply, output gap, YoY real wage change, unemployment rate and RER, respectively. Correspondingly, $\Delta_{f,k}^{h}$ is the parameter of variable f in equation k for lag h and v_t^f is the prediction error of variable f in time period t.

In order to select the order p for each country, I use Akaike's Information Criterion (AIC) since it balances under-fitting and overfitting by selecting the value p which minimizes the criterion (Snipes and Taylor 2014). Additionally, Ivanov and Kilian (2005) argue that AIC tends to create more accurate structural and semi-structural impulse-response compared to Hannan-Quinn Criterion (HQC) and Schwarz Information Criteria (SIC)[18]. The time periods where divided into estimation and testing data as in the benchmark (Table 1). Below, in table 3, are the values of p for the VAR(p) model of each country.

Table 3: VAR order for each country

Country	AIC	VAR(p)
United States	3	VAR(3)
Euro Area	3	VAR(3)
Netherlands	2	VAR(2)
Chile	4	VAR(4)
Mexico	12	VAR(12)

Once estimated the model's parameters with the 'vars' package by Pfaff (2008) , I produce a 12 months forecast for each country. Notice that United States' forecasted inflation is in the FED's target inflation rate, while future inflation for the Euro Area is leaning towards the lower bound of ECB's target inflation. Forecasted inflation in the Netherlands, however, seems to be more uncertain, since the 95% confidence interval[19] goes beyond 4% in the upper bound and reaches deflationary territory in the lower bound. The situation of Chile's inflation forecast is even more worrisome, since uncertainty is even greater than in the dutch case. The 95% confidence interval goes beyond 6% and below 0%, while the 95% confidence interval ranges from -2% inflation to 8%. The Mexican case is more optimistic, inflation ranges from 2% to 4.5% with a 80% confidence, while the 95% confidence interval ranges from 1% to above 5%.

[18] Although, HQC performs better for quarterly data with more than 120 samples, while SIC performs better for quarterly data with less than 120 observations.

[19] The 95% confidence interval is represented by the light blue fan and the 80% confidence interval by the dark blue fan.

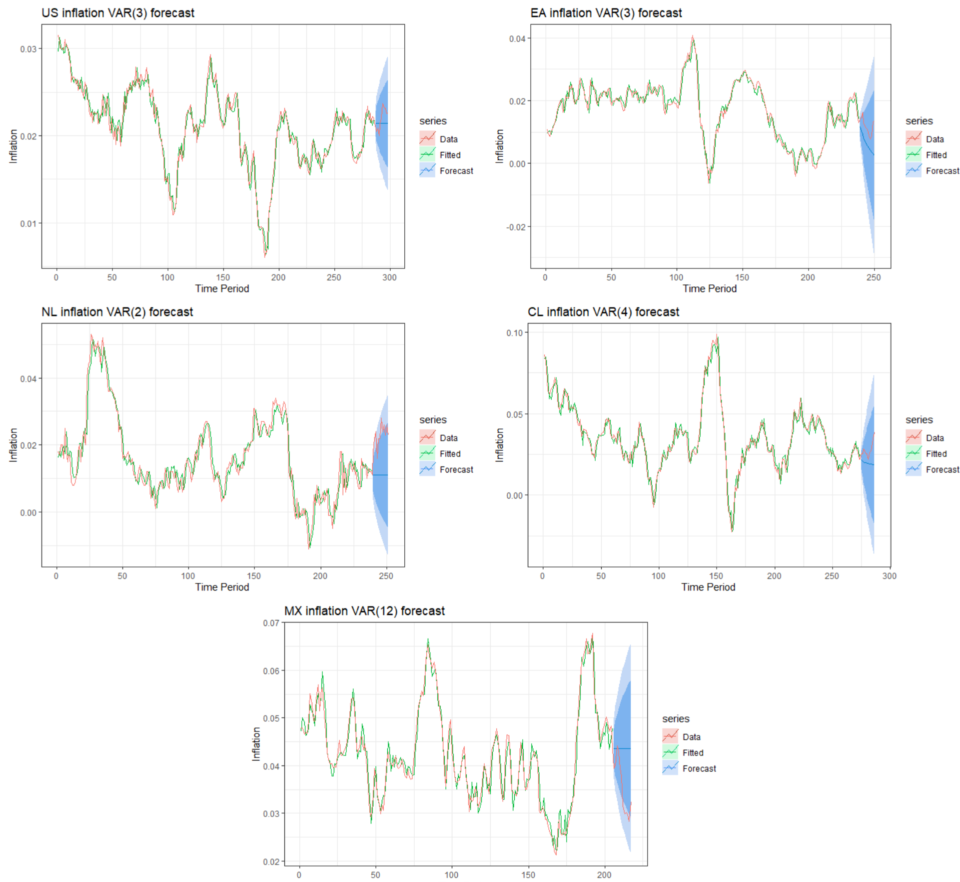

Figure 6: Inflation forecasts from a VAR model for different countries. Forecast fans are for confidence intervals of 80% and 95%

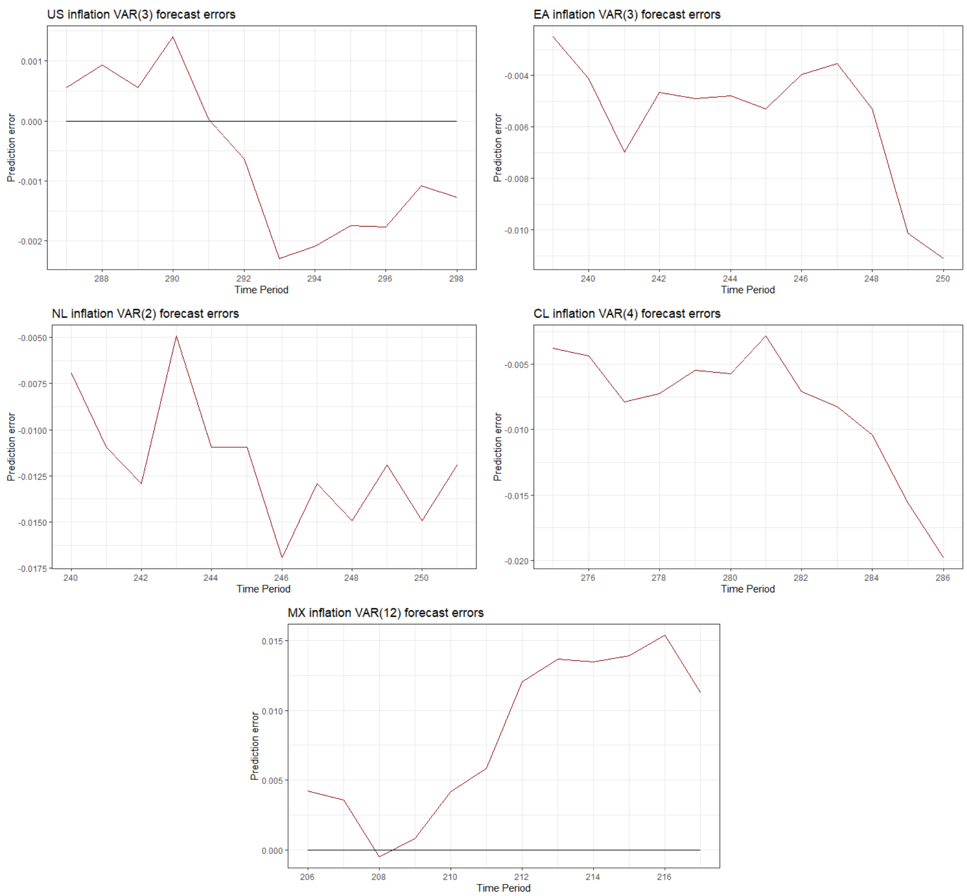

Figure 7: Inflation forecasts errors from a VAR model for different countries

The model overestimates inflation in the United States in the first months and underestimates it from June 2019 (time period 291) onward. However, the model tends to underestimate future inflation for all the forecast in the Euro Area, the Netherlands and Chile. Finally, the VAR model for the Mexican economy overestimates inflation during most of the forecast, being April 2019 (time period 208) the exception.

2.3 Neural Network

The use of traditional econometric methods such as simple autoregressive, ARIMA and VAR models are dominant in economic literature. Nonetheless, these type of models are limited to the completeness of data and the assumption of linear relationships between variables (Brownlee 2018). Machine learning algorithms do not require assumptions about data, thus allowing for more reliable out-of-sample predictions when overfitting is prevented[20].

Among different machine learning algorithms, Recurrent Neural Networks incorporate sequences to the model, i.e. the model learns from the relation between inputs and outputs over time respecting the order of observations (Brownlee 2018). Additionally, these types of neural networks learn temporal dependence found on data, meaning that the network can take observations and learn what previous data is relevant for producing a forecast.

Among the advantages mentioned by Brownlee (2018) of neural networks are (i) the ability to learn unspecified complex functions for mapping inputs to outputs, (ii) the lack of a scaling and stationary requirement of the data to produce a good model[21] and (iii) the possibility to easily make hybrid models with characteristics of multilayer perceptrons, convolutional and recurrent neural networks, exploiting the advantages of each type of deep learning algorithm.

As mentioned above, I will build a multivariate multi-step LSTM model for inflation forecasting. The main reason for using LSTM hidden layers instead of gated recurrent units (GRU) is that the former is more established in the literature when dealing with long-term dependencies[22] (Chung et al. 2014), however, choosing between LSTM and GRU units is a matter of taste since there is no conclusive evidence of one type of recurrent neural network outperforming the other (Chung et al. 2014, Petneházi 2019). LSTM layers contain three gates containing sigmoid activation functions to filter information from data. When inputs enter an LSTM layer, they are combined with previous information from the network's hidden state, then this combined data passes through the forget gate, which select what data should be forgotten[23]. Then, data travels to the input layer to be updated through a predetermined activation function and is weighted between zero and one according to its importance for prediction. Finally, the output gate decides what the new hidden state should be by updating the layer's output with the previous hidden state. This way, LSTM cells are connected to cells in the next layer as well as themselves.

[20] Additionally, it is technically easier to incorporate prior knowledge into traditional econometric models than into machine learning algorithms.
[21] Scaling and stationarity may not be required, but can improve the performance of neural networks.
[22] Long-term dependencies mean that information provided long time in the past can help provide context and predict.
[23] If the sigmoid activation function yields zero, the information is forgotten by the network, the closer it is to one, the better it remembers it

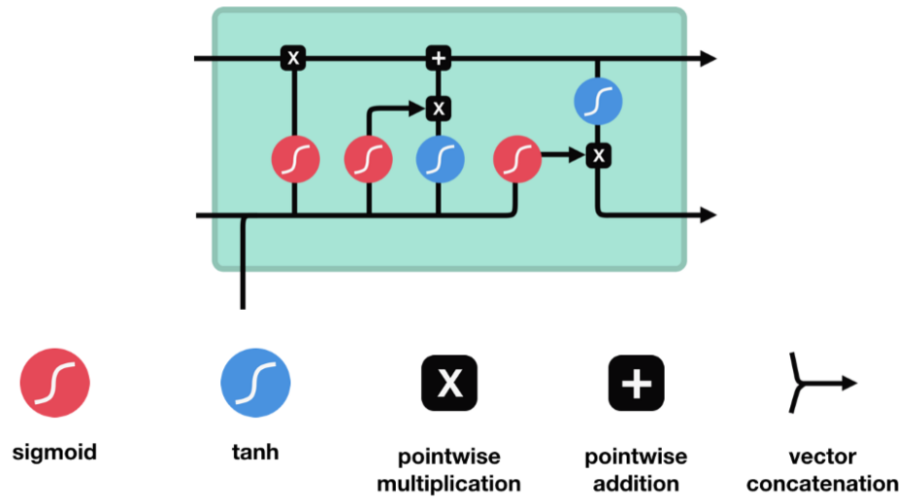

Figure 8: Internal Mechanism of an LSTM cell. Inputs are information from the previous layer and from the hidden state. The output of the LSTM cell is a combination of new processed information and the cell's memory, and the update of the hidden state

Source: https://towardsdatascience.com

2.3.1 Hyperparameters

Hyperparameters are those in charge of the learning process of a neural network thus they must be set beforehand by the practitioner. The choice of hyperparameters is usually performed using rules of thumb (Claesen and De Moor 2015) such as error in the validation set compared to training set and its slope, not too flat or too steep. I separated hyperparameters between individual and shared across countries in the sample.

The hyperparameters of the LSTM neural network, presented in this paper, shared by all countries in the sample are the network's structure,i.e. number of layers and neurons per layer, the activation function used and the early stopping algorithm. The model consist of a total of six layers[24], (i) the input layer containing seven[25] variables with twelve lags in order to use a year of past data to predict variables, (ii) four hidden layers, two LSTM layers with 100 neurons each and two Dense layers with 50 and 25 layers respectively. The output layer has twelve neurons, one per month of the forecast. The activation function was set as the hyperbolic tangent, since its output ranges from negative to positive one. Finally, I set the early stopping algorithm to minimize the loss of the validation set with a minimum improvement of 0.01 per epoch and a patience of 30 epochs, that is, training will stop after 30 iterations with no improvement.

[24] Šestanović (2019) and Choudhary and Haider (2012) mention that three layers are the minimum.
[25] Due to the lack of a long enough time series for real wages, the Euro Area's input layer contains only six lagged variables.

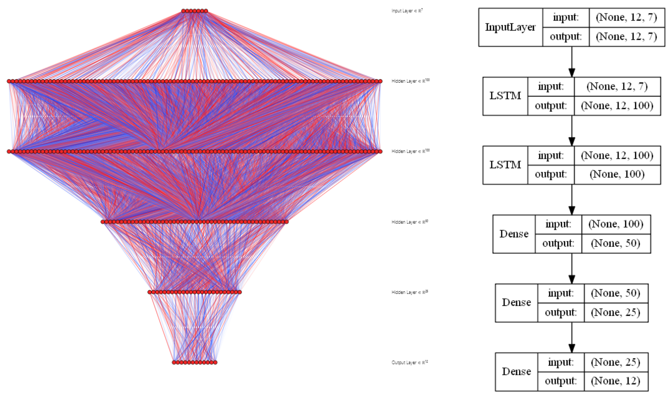

Figure 9: Model's simplified visualization (left) and model's structure (right)

Regarding individual hyperparameters, I set the dropout[26] and learning rates such that the validation learning curve lower than the training curve and have an appropriate slope, in other words, by checking the model does not have an overfitting problem. Using this rule of thumb I set dropout and learning rates to 0.4 and 0.0001 for the United States, 0.7 and 0.04 for the Euro Area as a whole, 0.3 and 0.02 for the Netherlands, 0.6 and 0.000025 for Chile and 0.8 and 0.0001 for Mexico.

Table 4: Individual Hyperparameters. Dropout includes recurrent dropout

Country	Dropout	Learning rate
United States	0.4	0.000100
Euro Area	0.7	0.040000
Netherlands	0.4	0.010000
Chile	0.6	0.000025
Mexico	0.8	0.000100

2.3.2 Training and results from the LSTM Neural Network

In order to train and validate the LSTM neural network for each country, I divide the data in the following manner. Using data from the beginning until January 2015 for training, from January 2015 to January 2019 for validation, leaving the February 2019 to January 2020 period for testing with out-of-sample predictions. This allows all validation sets to be the same length.

[26] I set dropout and recurrent dropout the same.

Table 5: Data splitting for neural network

Country	Training	Validation	Testing
United States	Jan 1995 - Jan 2015	Jan 2015 - Jan 2019	Feb 2019 - Jan 2020
Euro Area	Jan 1999 - Jan 2015	Jan 2015 - Jan 2019	Feb 2019 - Jan 2020
Netherlands	Jan 1999 - Jan 2015	Jan 2015 - Jan 2019	Feb 2019 - Jan 2020
Chile	Jan 1996 - Jan 2015	Jan 2015 - Jan 2019	Feb 2019 - Jan 2020
Mexico	Jan 2001 - Jan 2015	Jan 2015 - Jan 2019	Feb 2019 - Jan 2020

One problem of neural networks is they only provide on-point predictions. To address this issue, I train 50 LSTM networks from scratch to produce 50 different forecasts using the 'keras' package for Python per country due to the stochastic nature of neural networks. Using the different forecasts, I'm able to calculate a mean forecast and its standard deviation, generating more stable results and allowing the construction of a 95% confidence interval.

When training the network for the United States there were no signs of overfitting and there was convergence of the training process around the fifth epoch. Regarding inflation forecasting, the model manages to produce an accurate prediction of future inflation, on average underestimating inflation by 6 basis points (bps) with an error that ranges from -0.002 to 0.0015. It is important to note that observed inflation is always within the 95% confidence interval and that the interval ranges from 0.017 to 0.026, resulting in a much narrower error margin than in the benchmark and VAR models[27]. Additionally, the forecast predicts an average inflation of 2.14%, which falls within the target inflation range establised by the FED.

[27]The Benchmark and Var models' 95% confidence interval ranges from 0.015 to 0.027 and from 0.014 to 0.027, respectively

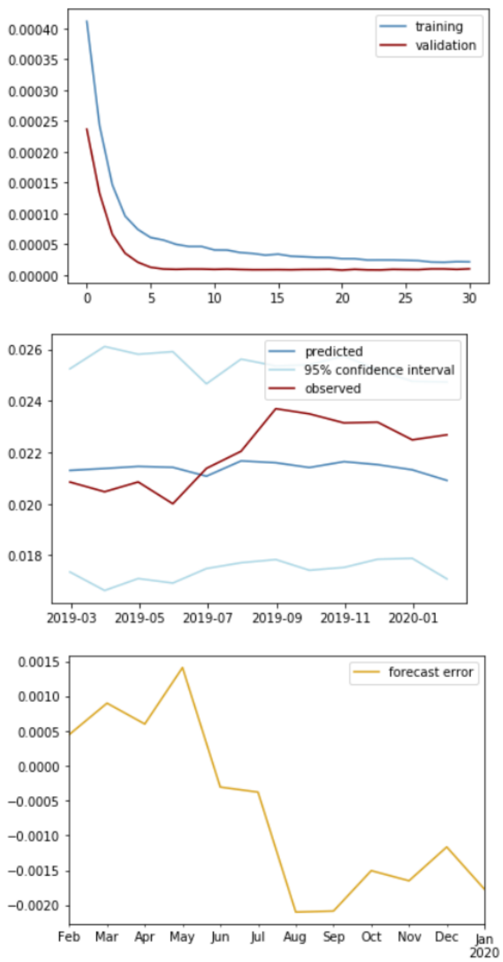

Figure 10: Results for the United States: (i) testing and validation curves (top), (ii) inflation forecast (middle) and (iii) errors from out-of-sample prediction (below)

In the case of the Euro Area as a whole, initial overfitting problems were solved by increasing the dropout and learning rates to 0.7 and 0.04. Although the dropout rate implies that the probability of a weight[28] being set to zero by the neural network is 70%, which may be considered as too high, there are precedents of neural networks using a dropout rate of 0.9 (Karpathy et al. 2014, Simonyan and Zisserman 2014, Varol and Schmid 2017). Regarding accuracy, the model predicts future inflation within its confidence interval during the whole forecasted period. However, it overestimates inflation by 72 bps on average, ranging from 0.004 to 0.0016. The confidence interval offers an error window of 240 bps on average, while the benchmark has a window of 250 bps and the VAR offers less certainty with 300 bps. Note that average forecasted inflation is consistent with ECB's target of close to but below 2%.

[28]Weights are referred to as parameters in traditional econometric models.

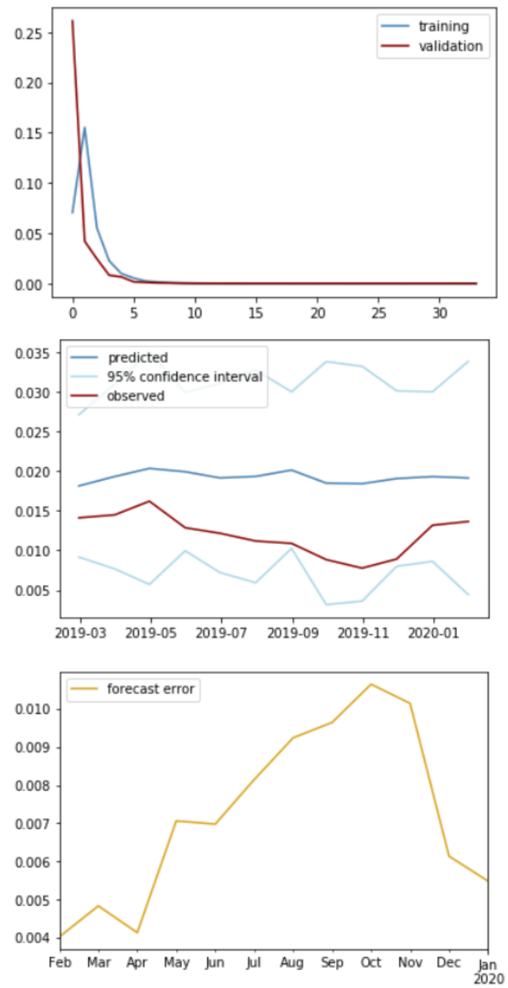

Figure 11: Results for the Euro Area: (i) testing and validation curves (top), (ii) inflation forecast (middle) and (iii) errors from out-of-sample prediction (below)

When applying the LSTM neural network to the Netherlans I had to set the dropout and learning rates to 0.7 and 0.2 respectively for the model to not overfit and meet the requirements to abide by the rules of thumb. Note that convergence is reached in the third epoch. Leaving technicalities behind, the model underestimates inflation by 75 bps on average, ranging from -0.012 to -0.002, and does not predict all future inflation within it confidence interval. A possible reason for this is the use of data on aggregated monetary policy instrument for a single country without adjusting by distribution, money is not supplied equally among the Euro Area members. Forecasted inflation is below observed inflation by 75 bps on average and the forecasted inflation rate is 1.53%, which is in line with ECB's target.

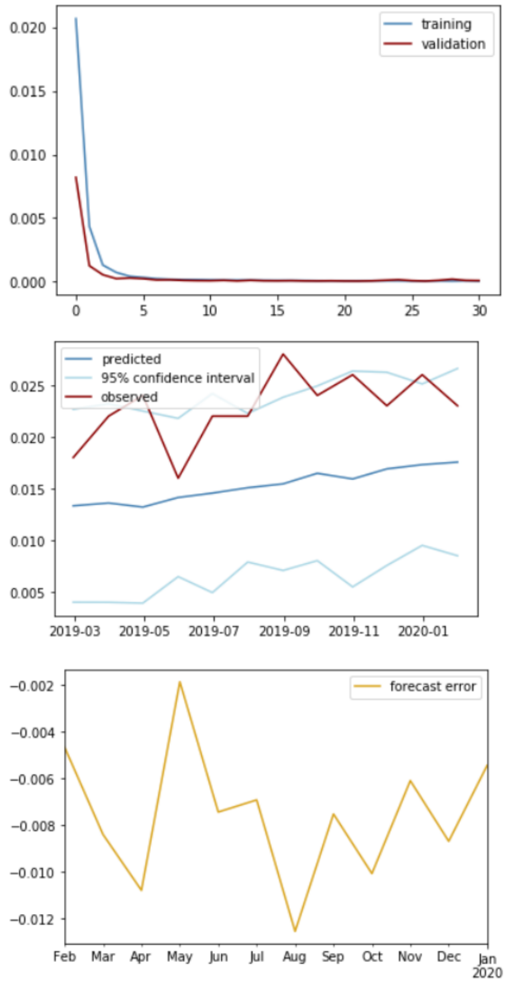

Figure 12: Results for the Netherlands: (i) testing and validation curves (top), (ii) inflation forecast (middle) and (iii) errors from out-of-sample prediction (below)

When adapting the hyperparameters to chilean data, I had to set the learning rate of the network to 0.000025, being by far the smallest, and dropout to 0.6. However, this yielded positive results, since is the training and validation curves clearly show a lack of overfitting and a good slope[29]. This results in a model that manages to replicate inflation for the most part and only fails to predict the last month's inflation within it 95% confidence interval. The model offers a margin of about 200 bps, while the VAR has a window between 400 and 900 bps, and the benchamrk's window is around 300 and 500 bps. The inflation forecast is in line with Banco Central de Chile's inflation target rate.

[29]If the learning rate is too high the slope becomes too steep and learning rates that are too low result in a curve that is too flat.

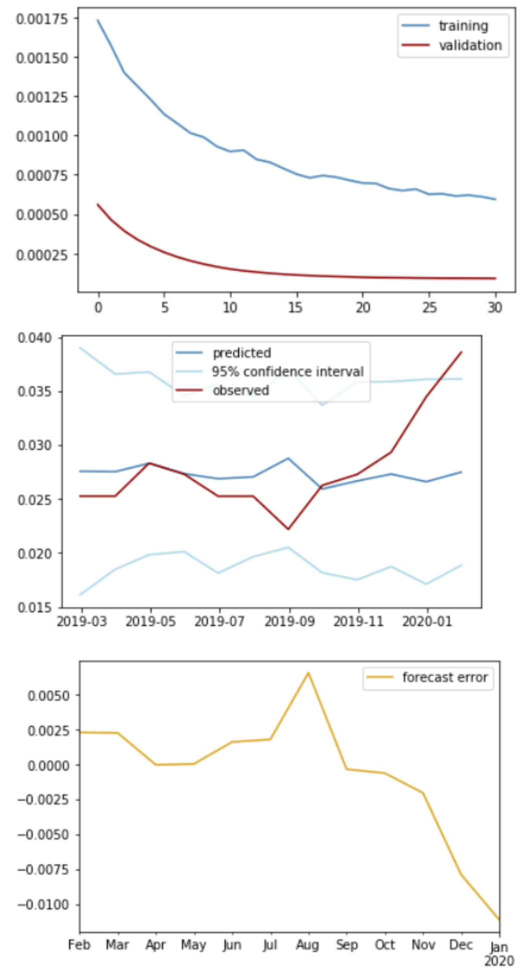

Figure 13: Results for Chile: (i) testing and validation curves (top), (ii) inflation forecast (middle) and (iii) errors from out-of-sample prediction (below)

Finally, the neural network was adapted to Mexico using dropout and learning rates of 0.8 and 0.0001. Predicted inflation is close to the upper bound of Banco de Mexico's target inflation and the model only manages to predict half of the forecasted period within its 95% confidence interval. On average, the neural network overestimates inflation by 15 bps, however this may not be very indicative for the mexican economy due to larger inflation variance[30].

[30]Observed inflation ranges from 2.8% to 4.4% while average predicted inflation is 3.7%.

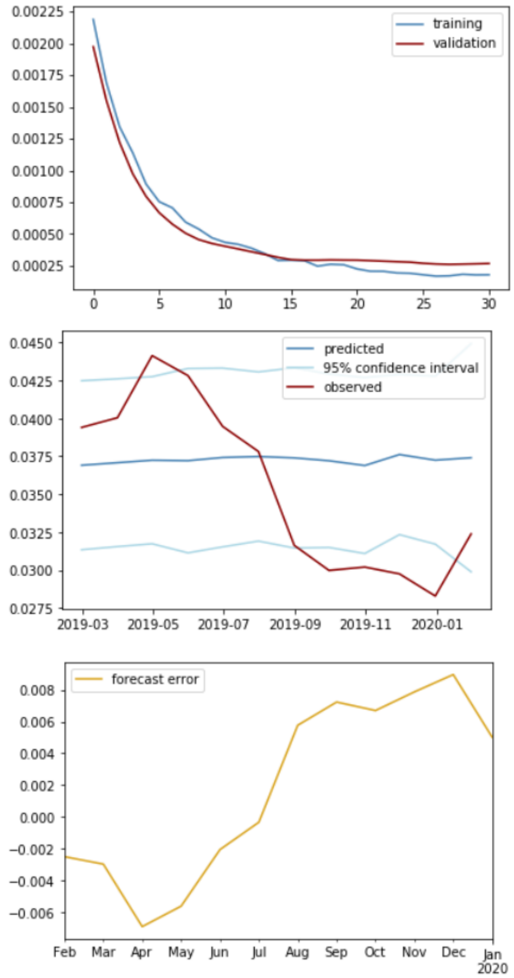

Figure 14: Results for Mexico: (i) testing and validation curves (top), (ii) inflation forecast (middle) and (iii) errors from out-of-sample prediction (below)

3 Models' Performance and Comparison

In this section I compare the performance of the different forecasting models in each country. To do so, I present three different metrics computed with RStudio: (i) mean absolute error (MAE), (ii) present bias average (PBA) and (iii) present punishing average (PPA). The first metric consist in a simple arithmetic average of the absolute value of the forecast errors and shows how good is a model at out-of-sample predictions overall. However, given that central banks and policy makers may be more interested in short term inflation, this metric may not be the optimal choice, because if a model produces large errors in long-run forecasts, while being accurate in the short-run, then the arithmetic average may be biased.

The second metric is a present biased average. The idea behind this performance measure is that central banks care more about short-run inflation than the long-run. To capture this, I use time diminishing weights to reduce the impact of later forecast errors. Therefore, the present biased average is computed as follows:

$$PBA^i = \frac{\sum_{t=1}^{T}(1 - \frac{t-1}{T})|F_t^i|}{T}, \tag{5}$$

where PBA^i is the present bias average of model i, t is the number of time periods in the future, T is the forecast horizon and F_t^i is the forecast error in t of model i.

Finally, I use the opposite of the PBA in order to assess long-run forecast accuracy. The PPA punishes immediate inflation by using weights to reduce the effect of short-run inflation while increasing the impact with time, biasing the metric towards the future. The equation for this metric is presented below.

$$PPA^i = \frac{\sum_{t=1}^{T}(1 - \frac{T-t}{T})|F_t^i|}{T}, \tag{6}$$

where PPA^i is the present punishing average of model i, t is the number of time periods in the future, T is the forecast horizon and F_t^i is the forecast error in t of model i.

When comparing mean average error across countries, the first thing we may note is that error is relatively low in most cases, being few the cases with an average error over 100 bps. Using the average of forecast errors metric, I find that in the case of the Netherlands, Mexico and Chile the LSTM neural network produces the lesser average error. In the United States all models perform similarly. I also find that the AR(1) process yields the lowest average error in the Euro Area.

In the case of present biased average, the results for the United States are the same across all models and this metric confirms the results to the mean absolute average metric. Finally, using the present punishing average metric, I find similar results to overall average and PBA. However, the difference in long-term accuracy of the LSTM is greater than short-term accuracy in the case of Mexico.

In general, the LSTM neural network has a greater time consistent accuracy in emerging market economies[31], meaning that accuracy in short-term predictions is similar to that of long-term predictions. This is noted by comparing Chile and Mexico's PBA and PPA. In the case of Chile, the VAR's errors increase from 63 bps to 101 bps when comparing PBA and PPA respectively. The case of Mexico is more extreme, since both the benchmark and the VAR's accuracy is time inconsistent, while the LSTM neural network only suffers an increase of 12 bps when going from short-term to long-term predictions.

[31] In the sample, Chile and Mexico.

Table 6: Results from mean absolute error, present biased average and present punished average

Country	Model	MAE	PBA	PPA
United States	Benchmark	0.0012	0.0010	0.0013
	VAR	0.0012	0.0010	0.0014
	LSTM	0.0012	0.0010	0.0014
Euro Area	Benchmark	0.0024	0.0021	0.0026
	VAR	0.0056	0.0048	0.0064
	LSTM	0.0072	0.0066	0.0078
Netherlands	Benchmark	0.0105	0.0099	0.0111
	VAR	0.0118	0.0108	0.0127
	LSTM	0.0075	0.0074	0.0077
Chile	Benchmark	0.0037	0.0028	0.0046
	VAR	0.0081	0.0063	0.0101
	LSTM	0.0031	0.0020	0.0041
Mexico	Benchmark	0.0090	0.0063	0.0116
	VAR	0.0082	0.0058	0.0106
	LSTM	0.0050	0.0044	0.0056

4 Conclusion and Final Remarks

The use of neural networks for inflation forecasting is a field that has not seen a big development when compared to other fields such as finance, one reason could be the lack of huge data sets for inflation since it is reported on a monthly basis while stock prices are available daily and even intraday. Other reason may be the idea of "if it is not broken, don't fix it" that surrounds central banks. However, this paper sums to the literature that explores the use of neural networks for Inflation forecasting.

I find that in two out of five countries, the use of an LSTM neural network provide a more accurate forecast than a VAR and an AR(1) model, while not being different in other three countries and only outperformed in one, according to the mean absolute error metric. These results are confirmed by the PBA and PPA metrics.

Given neural networks ability to capture non-linear relations in data, we may expect that changing the forecasted period to a less stable moment in time, such as 2001 or 2007 may show a better performance of the LSTM neural network compared to the AR(1) and VAR models. This seems to be the case of Mexico, a country with high inflation volatility, since the different average metrics show a superior performance of the LSTM neural network. However, linear models such as the benchmark and VAR may provide a better forecast when inflation has been relatively stable. This is seen in the case of the Netherlands, where inflation previous to the forecast period did not had high inflation volatility. It is important to remark that comparison results were obtained by comparing three single multi-step forecasts of length 12, and if the test window is changed from the February 2019 to January 2020 period to a different one, results may vary.

If we use the neural network's forecasted inflation as a proxy of expected inflation, then we would be able to measure the confidence of the general public in their central bank. This way, if predicted inflation deviates too much from the central bank's target, it could mean, under normal circumstances, that the central bank is not very effective controlling inflation or that it has other priorities. Following this idea, most countries had a predicted inflation close to their target inflation, the exception being Mexico with a predicted inflation closer to Banco de Mexico's target inflation upper bound rather than the on-point target.

Other research projects regarding inflation forecasting using neural networks may incorporate the evolution of commodity prices and the performance of financial indices as a measure of investment, since an increase in the yield of financial assets may be linked with inflationary pressures. It may also be interesting to compare the performance of GRU and LSTM neural networks in this context.

References

[1] Algorithmia. (2019). *2020 state of enterprise machine learning*. Retrieved from `https://info.algorithmia.com/hubfs/2019/Whitepapers/The-State-of-Enterprise-ML-2020/Algorithmia2020StateofEnterpriseML. pdf?hsLang=en-us.`

[2] Al-Maqaleh, B. M., Al-Mansoub, A. A., Al-Badani, F. N. (2016). Forecasting using artificial neural network and statistics models. *International Journal of Education and Management Engineering, 6*(3), 20-32.

[3] Allsopp, C., Kara, A., Nelson, E. (2006). United Kingdom Inflation Targeting and the Exchange Rate. *The Economic Journal, 116*(512), F232-F244.

[4] Andersson, M., Masuch, K., Schiffbauer, M. (2009). Determinants of inflation and price level differentials across the euro area countries.

[5] Athey, S., Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives, 31*(2), 3-32.

[6] Benigno, P., Rodríguez-Palenzuela, D., Altissimo, F. (2005). Long-run determinants of inflation differentials in a monetary union. *Moneda y crédito*, (220), 203-260.

[7] Bhandare, A., Bhide, M., Gokhale, P., Chandavarkar, R. (2016). Applications of convolutional neural networks. *International Journal of Computer Science and Information Technologies, 7*(5), 2206-2215.

[8] Binner, J. M., Elger, C. T., Nilsson, B., Tepper, J. A. (2006). Predictable non-linearities in US inflation. *Economics Letters, 93*(3), 323-328.

[9] Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

[10] Blanchard, O. J. B, A., Giavazzi, F., Amighini. (2010). Macroeconomics: A European Perspective. Financial Times Prentice Hall.

[11] Boelaert, J., Ollion, É. (2018). The Great Regression. *Revue française de sociologie, 59*(3), 475-506.

[12] Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science, 16*(3), 199-231.

[13] Brownlee, J. (2018). *Deep learning for time series forecasting: Predict the future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery.

[14] Chakraborty, C., Joseph, A. (2017). *Machine learning at central banks* (No. 674). Bank of England.

[15] Choudhary, M. A., Haider, A. (2012). Neural network models for inflation forecasting: an appraisal. *Applied Economics, 44*(20), 2631-2635.

[16] Chung, J., Gulcehre, C., Cho, K., Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

[17] Claesen, M., De Moor, B. (2015). Hyperparameter search in machine learning. *arXiv preprint arXiv:1502.02127*.

[18] Cuche-Curti, N. A. (2006). Policy-relevant models for central banks. *Swiss National Bank Quarterly Bulletin, 2*, 54-63.

[19] Einav, L., Levin, J. (2014). Economics in the age of big data. *Science, 346*(6210), 1243089.

[20] Faria, J. R., Carneiro, F. G. (2001). Does high inflation affect growth in the long and short run?. *Journal of applied economics, 4*(1), 89-105.

[21] Gardner, M. W., Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment, 32*(14-15), 2627-2636.

[22] Greene, W. H. (1998). *Econometric analysis 3th edition*. International edition, New Jersey: Prentice Hall.

[23] Groen, J. J., Paap, R., Ravazzolo, F. (2013). Real-time inflation forecasting in a changing world. *Journal of Business Economic Statistics, 31*(1), 29-44.

[24] Hartmann, P., Smets, F. (2019). The First 20 Years of the European Central Bank: Monetary Policy.

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

[26] Hyndman, R. J., Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software, 26*(3), 1–22.

[27] Ivanov, V., Kilian, L. (2005). A practitioner's guide to lag order selection for VAR impulse response analysis. *Studies in Nonlinear Dynamics Econometrics, 9*(1).

[28] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 1725-1732).

[29] Kvint, V. L. (2004). *The Global Emerging Market in Transition: Articles, Forecasts, and Studies, 1973-2003*. Fordham Univ Press.

[30] Lim, Y. C., Sek, S. K. (2015). An examination on the determinants of inflation. *Journal of Economics, Business and Management, 3*(7), 678-682.

[31] Morsy, H., Jaumotte, M. F. (2012). *Determinants of inflation in the euro area: The role of labor and product market institutions* (No. 12-37). International Monetary Fund.

[32] Nakamura, E. (2005). Inflation forecasting using a neural network. *Economics Letters, 86*(3), 373-378.

[33] Petneházi, G. (2019). Recurrent neural networks for time series forecasting. *arXiv preprint arXiv:1901.00069*.

[34] Pfaff, B. (2008). VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software 27*(4).

[35] Pirovano, M., Van Poeck, A. (2011). *Eurozone inflation differentials and the ECB*. University of Antwerp, Faculty of Business and Economics.

[36] Poole, W. (1970). Optimal choice of monetary policy instruments in a simple stochastic macro model. *The Quarterly Journal of Economics, 84*(2), 197-216.

[37] Šestanović, T. (2019). Jordan neural network for inflation forecasting. *Croatian Operational Research Review*, 23-33.

[38] Sherstinsky, A. (2020). Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. *Physica D: Nonlinear Phenomena, 404*, 132306.

[39] Simonyan, K., Zisserman, A. (2014). Two-stream convolutional networks for action recognition in videos. In *Advances in neural information processing systems* (pp. 568-576).

[40] Sirignano, J., Spiliopoulos, K. (2018). Mean field analysis of neural networks. *arXiv preprint arXiv:1805.01053.*

[41] Snipes, M., Taylor, D. C. (2014). Model selection and Akaike Information Criteria: An example from wine ratings and prices. *Wine Economics and Policy, 3*(1), 3–9.

[42] Teräsvirta, T., Van Dijk, D., Medeiros, M. C. (2005). Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting, 21*(4), 755-774.

[43] Varol, G., Laptev, I., Schmid, C. (2017). Long-term temporal convolutions for action recognition. *IEEE transactions on pattern analysis and machine intelligence, 40*(6), 1510-1517.

[44] Walsh, C. E. (2010). *Monetary theory and policy*. MIT press.

[45] Wibisono, O., Ari, H. D., Widjanarti, A., Zulen, A. A., Tissot, B. (2019). The use of big data analytics and artificial intelligence in central banking. *The Capco Institute Journal of Financial Transformation*, Data analytics, 50th edition, 70-83.