Erasmus University Rotterdam
Eramsus School of Economics

Bachelor Thesis Econometrics and Operations Research
Quantitive Finance

Estimating Nonlinear Relation between Fiscal Policy and Macroeconomic Activity using Quantile Regressions and Neural Networks with Vector Autoregressive Models

Abstract
We estimate the relation between fiscal policy and macroeconomic activity, where we use output as a measure for the macroeconomic activity. To do so, we make use of two nonlinear methods, namely quantile regressions and neural networks. Both methods are performed in combination with vector autoregressive (VAR) models. Also, we evaluate the forecasting performance by using an expanding window. We find that the neural networks outperform the quantile regressions in both, estimation and prediction, of the nonlinear relation.

Name student: **Amber Ilyas**
Student ID number: **466757**

Supervisor: **Dr. A.M. Schnucker**
Second assessor: **Dr. Y. Sun**

Date final version: **19 July 2020**

Contents

1 Introduction

The relation between government spending and output is an interesting topic for many policy makers. Estimating the precise relation between these two variables is rather difficult, since standard linear regressions capture only the central tendency of effects (Linnemann & Winkler, 2016). Many papers suggest that the effect of government spending shocks on US macroeconomic activity is not linear. Makinng use of nonlinear methods is then a more suitable approach. Therefore, this paper estimates the relation between fiscal policy and economic activity, using two different nonlinear methods; quantile regressions and neural networks. Both methods are used in combination with a vector autoregression (VAR) model.

Linear models focus only on the mean of the conditional distribution of the dependent variable. Quantile methods on the other hand, have the advantage that they are able to examine how the whole conditional distribution of the outcome variable gets affected by changes in the explanatory variables. The effects may differ at various parts of the distribution. Linnemann & Winkler (2016) perform quantile regressions to evaluate whether output (GDP) responds different to changes in fiscal policy, when the conditional distribution of output is in its lowest or highest 10 per cent.

Nowadays, there are many more methods in order to estimate and predict nonlinear relations. Nonlinear machine learning methods seem to have interesting features, which make it easier to catch the nonlinear trend. To estimate the relation between government spending and macroeconomic activity, we use another nonlinear method, next to quantile regressions, namely a neural network (Hornik, Stinchcombe, & White, 1989); (Gu, Kelly, & Xiu, 2020). This paper deals with the following research question: *"Does a neural network outperform a quantile regression in terms of estimation and prediction of the nonlinear relation between US government spending and US output?"*

According to (Linnemann & Winkler, 2016), quantile regressions have advantages, both relative to linear methods and nonlinear methods. Quantile methods do not depend on the regime the economy inhabits at. In this case, these regimes are the recessions and booms. Furthermore, it is not necessary to choose an index as a measure of the business cycle. Also, the quantile regressions are capable of estimating the impact of changes in explanatory variables on the whole distribution of the outcome variable, since there are no restrictions for the parameters. Another advantage of this method is that it does not demand an exact definition of the regimes. Moreover, quantile regressions are able to pick up the nonlinearities that can not be estimated, since the economy is in an expansion or recession.

Linnemann & Winkler (2016) also mention some disadvantages of the quantile regressions. Quantile methods do not present a specific and testable model which investigates the nonlinear behaviour in the data. Also, policy makers might experience it as a challenge to use the results of a quantile regression, since they do not know the exact quantile of macroeconomic activity that the economy resides in. However, they have an indication whether the economy is in a lower or higher quantile than its mean/median.

Machine learning is a more modern approach. Gu, Kelly and Xiu (2020) state in their paper that a neural network is one of the most powerful modeling device in machine learning. This method is known for its flexibilty. A neural network may outperform a quantile regression in terms of estimation and prediction of the nonlinear relation between fiscal policy and output. As a result, policy makers can use this nonlinear relation to predict the macroeconomic activity.

To determine whether neural networks outperform quantile regressions in estimation and prediction of the nonlinear relation between government spending and output, we first reproduce a part of the results of the paper (Linnemann & Winkler, 2016). This is the first model. For the second model, we use a fully connected feed-forward neural network.

To estimate and predict the relation between US output and US government spending, we use a data set which contains quarterly US data from 1966Q1 to 2013Q4. The variables *Government Spending, Output, Net Taxes* and *Interest Rate* are included in the VAR model. We find that the neural network estimates a larger proportion of the relation between government spending and output, than the quantile regressions. Moreover, it predicts the relation better; the MSFE is lower and the out-of-sample R^2 is higher compared to the MSFE and out-of-sample R^2 of the quantile regressions.

In the rest of the paper we first discuss relevant literature in the literature section. Thereafter, we describe the variables used for the VAR models in the data section. We then elaborate on the methods used for both models to estimate the relation between government spending and output. Next, we explain how the impulse responses are constructed in order to obtain how output and government spending respond to a fiscal expansion. The measures used to evaluate the predictive performance of these models are also discussed in this section. Subsequently, the result section presents the results of the two models. After this, we answer our research question by making a comparison between the two models in the conlusion section. Finally, the discussion mentions the limitations of the methods used for this research.

2 Literature

Literature proposes various methods to estimate the relation between fiscal policy and US macroecomic activity. Both linear and nonlinear methods have been used.

Blanchard & Perotti (2002) made use of linear vector autoregressive models in combination with event studies to estimate the aforementioned relation. The results of (Blanchard & Perotti, 2002) indicate that an increase in government spending leads to an increase in output as well, but output decreases when there is a positive change in tax. However, linear models are not able to estimate the nonlinear impact of a fiscal expansion on the macroeconomic activity.

Other papers used nonlinear statistical methods. The paper (Auerbach & Gorodnichencko, 2012) investigates how the economy responds to a change in fiscal policy, by using "regime-switching models". After obtaining the results, they state that fiscal policy has a greater impact on the macroeconomic activity when the economy is in a recession rather than in an expansion. Blanchard & Leigh (2013) estimate how the forecast errors for the growth of the economy are related to the plans the government made in order to reduce their deficits and total debt, during the crisis. They find that governments of developed countries had a greater fiscal consolidation strategy when the growth of their economy was lower than predicted. According to (Blanchard & Leigh, 2013), predicting the results of fiscal consolidation too positively might lead to an underestimation of the size of the effects of changes in government spending and tax.

To obtain results for our first model, we will use the same method as described in (Linnemann & Winkler, 2016). They use quantile regressions in two different contexts. As measures for US macroeconomic activity, they either use output or the unemployment rate.

First they estimate vector autoregressions (VAR) models by performing quantile regressions. They also evaluate the impact of a fiscal expansion on output (GDP) by estimating the quantile-specific impulse responses. For the second framework, they evaluate the impact of government spending on predictions of output for different quantiles. To do so, they use the local projection method (Blanchard & Perotti, 2005). They find that a change in government spending affects the lowest decile of output more than higher deciles of this variable. On the other hand, for the highest quantile of unemployment, higher government spending decreases the unemployment in the US.

There has been numerous studies, that apply machine learning (ML) methods to solve complex finacial problems. The papers (Feng, Giglio, & Xiu, 2017) and (Kelly & Pruitt, 2015) use linear and nonlinear machine learning methods such as Partial Least Squares (PLS) and Neural Networks. These methods are able to handle a huge amount of explanatory variables. Gu et al. (2020) state that the use of machine learning methods in a financial context results in huge benefits for the economy. Machine learning methods are able to increase the amount of variables, without negatively affecting the R-squared of the model. Machine learning methods are very flexibily, which has a positive effect on the estimates of the model. The paper also states that a neural network is the preferred nonlinear method when it comes to machine learning.

3 Data

In this section the data set which is used to perform the quantile regression and to create a neural network is described. First, in section 3.1, the samples will be discussed. Section 3.2 shows the variables used for this paper and their descriptive statistics.

3.1 Calibration

The data set used for our research contains US data from 1966Q1 to 2013Q4, with 236 quarterly observations. To obtain the predictive performance of the models, we split our sample into sub-samples. For the quantile regressions, we use roughly 80 % of the data to estimate the model. This is the estimation sample, which spans 1955Q1-2002Q4. The remaining 20 % is predicted by means of an expanding window. This sub-sample is the test sample and contains 44 observations, from 2003Q1 till 2013Q4. For our second model, the neural network, we split the sample in three sub-samples: the estimation, validation and test sample, which contain 60, 20 and 20 per cent of the observations, respectively. The validation sample makes sure that the neural network is shielded from overfitting (Gu et al., 2020). The estimation sample spans 1955Q1-1990Q4, the validation 1991Q1-2002Q4 and for the test sample we still use the observations from 2003Q1 till 2013Q4.

3.2 Variables

The variables are the same as used in (Linnemann & Winkler, 2016). The data set is obtained from the FRED database. This database contains many economic time series from US and international sources. The data set used for this paper consists of the following four variables:

- **Government Spending**: real government consumption and gross investment
- **Output**: real gross domestic product (GDP)
- **Net Taxes**: measured as the real value of government current tax receipts deflated with the GDP deflator
- **Interest Rate**: the short-run real interest rate, constructed as the annualized difference between the Federal Funds Rate and the log-change

Linnemann & Winkler (2016) use another variable; the ratio of government debt held by the public. Since there was no quarterly data available for this variable, for the period 1955Q1 to 2013Q4, we did not include it.

All variables are selected in order to take influences into account, that might have an effect on the nonlinear relation between output and government spending. To detrend the data, we modify the variables government spending, output and net taxes by taking the differences between the log of these variables and the quadratic time trends. This makes it easier to compare our results with earlier literature, since most empirical studies measure their variables this way. Table 1 shows the descriptive statistics of the four variables. The range of the variable *Net Taxes* is larger than the range of the other three variables. That's why *Net Taxes* has the highest standard deviation among the four variables. Also, the mean of *Interest Rate* is higher than the mean of the other variables.

Table 1: Properties of the variables.

	Mean	Maximum	Minimum	Std. Dev.	Observations
Government Spending	-3.39E-08	0.141	-0.102	0.049	236
Output	1.20E-08	0.057	-0.087	0.033	236
Net Taxes	-8.47E-09	0.323	-0.441	0.141	236
Interest Rate	0.019	0.101	-0.027	0.026	236

Note: The table displays the descriptive statistics of the four variables.

4 Methodology

As we mentioned in the introduction, in this paper we estimate the nonlinear relation between government spending shocks and US macroeconomic activity, by using two nonlinear methods. We compare the two models to conclude which method is more suitable for this problem. In section 4.1, the quantile regressions, which are used for the first model, are described first. After that, the neural network is discussed. To estimate the effect of government spending shocks on output, we calculate the impulse responses by means of a QVAR (**Q**uantile **V**ector **A**uto**r**egressive) model. Section 4.2 covers this part. The methods used to evaluate the predictive power of the two models are explained in section 4.3 .

4.1 The two models

The two models used for this research are described below. Both models are used in combination with a VAR model.

4.1.1 Quantile Regressions

For the first model, the same method will be used as (Linnemann & Winkler, 2016) did in their paper. We will apply quantile regression estimation to a VAR model. This model consists of a constant and four lags of the four variables in the vector autoregressions (VAR). (Cecchetti & Li, 2008) have been the first to use quantile regressions in the context of VAR models.

Quantile regressions, as described by (Koenker & Bassett, 1987), are an extensive form of the basic regressions. To understand how quantile regressions work, it is required to be familiar with standard least squares regressions. These regressions estimate the conditional mean of the dependent variable y_t across values of a vector, x_t, which contains explanatory variables, such that

$$E(y_t|x_t) = x_t\beta \tag{1}$$

The model aims to minimize the sum of squared residuals, ϵ^2, by estimating the parameters for the explanatory variables x_t, which are presented by the vector β. The dimensions of x and β are (n x k) and (k x 1) respectively, where n denotes the amount of observations and k the amount of explanatory variables. This linear model explains variations in the mean of y_t, caused by changes in the explanatory variables, x_t.

Whereas the least suqares method only focuses on the conditional mean, quantile regessions take the whole distribution of the dependent variable into account. The quantiles, $q \in (0, 1)$, describe the distribution. When $q = 0.1$ for example, the output is in the lowest 10 percent of its conditional distribution. Similarly, if $q = 0.9$, output is in the highest 10 percent of its conditional distribution. The model then explains changes in this part of the distribution of the dependent variable as a result of variations in the explanatory variables. The quantile function $Q_q(.)$ for quantile q is $Q_q(y_t) = F^{-}1(q)$, where $F(y_t)$ is the PDF of the variable y_t. Given x_t, the qth quantile of the dependent variable is explained as follows:

$$Q_q(y_t|x_t) = x_t\beta(q) \tag{2}$$

Here, the vector β contains values which show how the dependent variable is affected by a change in the corresponding explanatory variable. The difference between standard least squares regressions and quantile regressions is visible in the notation β_q. For each quantile q, there is a different vector β. This vector is estimated as follows:

$$\hat{\beta}(q) = argmin_{\beta(q)} \sum_t \rho_q[y_t - x_t(\beta(q))] \tag{3}$$

This estimation is given by (Koenker & Bassett, 1987). Let $\omega_t = y_t - x_t(\beta(q))$, then $p_q[\omega_t] = (q - I_{\omega_{t<0}})\omega_t$. Here $I_{\omega_{t<0}}$ is an indicator function which is equal to one if $\omega_t < 0$ and zero otherwise.
The parametric vector $\hat{\beta}(q)$ is an estimation of the marginal effect of the explanatory variables. Since basic regressions only focus on changes in the conditional mean of y_t, the results we obtain by performing quantile regressions contain much more information. By estimating the model for different quantiles, the influence of x_t on the whole distribution of y_t can be evaluated, rather than only on the mean of y_t.

For this model, a VAR(4) model (VAR(p), with $p = 4$ lags) is used. Let $z_t = (z_{1t}, z_{2t}, ..., z_{kt})'$ be a vector of k variables measured at time t and let $q = (q_1, q_2, ..., q_k)'$ be a vector of k quantiles. In this case, k is equal to 4, since we are using four variables. For the quantile regressions, the order of the variables is very important, since the variables have to be in line with their corresponding quantiles. Our model uses the explanatory variables in the following order: **government spending** (g_t), **output** (y_t), **net taxes** (τ_t) and **interest rate** (r_t).

Since we are estimating the effect of government spending shocks on different quantiles of output (0.1, 0.5 and 0.9), we are using the same model thrice, with a different vector q. These vectors are:

$$q = \begin{pmatrix} 0.5 \\ 0.1 \\ 0.5 \\ 0.5 \end{pmatrix} \qquad q = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{pmatrix} \qquad q = \begin{pmatrix} 0.5 \\ 0.9 \\ 0.5 \\ 0.5 \end{pmatrix}$$

The linear QVAR model is then as follows:

$$Q_q(z_t|z_{t-1}, ..., z_{t-4}) = c(q) + \sum_{i=1}^{4} B_i(q) z_{t-i}, \tag{4}$$

where

$$B_i(q) = \begin{pmatrix} \beta_{i,11}(q_1) & ... & \beta_{i,14}(q_1) \\ \beta_{i,21}(q_2) & ... & \beta_{i,24}(q_2) \\ \beta_{i,31}(q_3) & ... & \beta_{i,34}(q_2) \\ \beta_{i,41}(q_4) & ... & \beta_{i,44}(q_4) \end{pmatrix}, \qquad and \quad c(q) = \begin{pmatrix} c_1(q_1) \\ c_2(q_2) \\ c_3(q_3) \\ c_4(q_4) \end{pmatrix}$$

Here, the coefficients $\beta_{ijn}(q_j)$ show how the qth quantile of the conditional distribution of variable z_{jt} gets affected by lag i of variable n. The constant c is also different for each quantile q. This is obvious from the notation c_q. Thus, the estimated model is able to determine how a change in the q_i^{th} quantile of the conditional distribution of z_{it} gets affected by the q_j^{th} quantile of the conditional distribution of the jth variable, z_{jt}.

VAR models have the advantage that the right-hand side of each equation is the same. In our case this side contains the four variables, with four lags for each variable. Thus, the model can be estimated equation-by-equation.

4.1.2 Neural Network

For our second model, we use a "fully connected feed-forward neural network". Gu et al (2020) states that a neural network is one of the most powerful machine learning techniques. Neural networks were also used in the paper (Akimov, Azagouag, Djibuti, Ilyas & Lingsveld, 2020). In this section, we elaborate on this machine learning method.

A neural network is able to handle difficult machine learning problems, such as computer vision and automated game-playing. Its most attractive feauture is its flexibility. This felixibility is caused by the large amount of hidden layers the networks can contain, which increases the performance of the model. However, the network also loses its transparency due to the hidden layers. It makes them extremely complex and difficult to interpret. Therefore, neural networks are also called "black boxes".

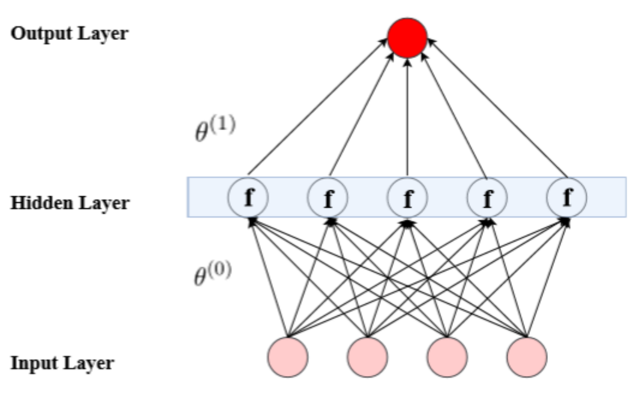

Figure 1: Neural Network, illustration from (Gu et al., 2020)

Neural networks contain different kinds of layers. The simplest neural network consists of an input layer and an output layer. See Figure 6 in the appendix. This linear model is identical to an ordinary least squares (OLS) regression. Let $x_1....x_k$ be our predictors and β a parametric vector. The dimension of the predictors is equal to the amount of units in the input layer. After giving the model its input, the network assigns weights to these variables and transfers them together with a constant to the output layer. Here, the weighted signals are aggregated into the forcast:

$$\beta_0 + \sum_{i=1}^{k} x_i \beta_i \tag{5}$$

The network becomes more flexible when we include hidden layers. The amount of hidden layers are noted as $l....L$ and the amount of neurons within each hidden layer are noted as $n....N^l$. These layers are added between the input and output layer. The variables from the input layer are passed on to the first hidden layer, together with a constant. Each hidden layer contains neurons and each neuron receives information from the input layer. The network then applies the activation function $f(x)$ to the neurons, which transforms the input variables nonlinearly. After that, signals are passed on to each neuron in the following hidden layer. These signals contain information from the previous layer. This goes on until the network reaches the last hidden layer, which then sends signals to the output layer. The output layer aggregates all the information from every neuron linearly as an target variable. To understand how this happens, we take a look at a neural network with only one hidden layer. The outcome variable is then computed as follows:

$$y(x, \beta) = \beta_0^{(1)} + \sum_{n=1}^{N^1} s_n^{(1)} \beta_n^{(1)} \tag{6}$$

Such a network is called a "fully connected feed-forward neural network": each neuron from each hidden layer receives information from each neuron from the previous hidden layer. An example of this network is shown in Figure 1. The pink circles denote the input layer and the red circle denotes the output layer. In between, there is one hidden layer, which contains 5 neurons. Each arrow between the different layers is associated with a weight parameter.

It is quite a challenge to set up the neural network. Many choices need to be made carefully, inlcuding the amount of parameters and hidden layers and the number of neurons within each hidden layer. Increasing this amount can ehance the performance of the model. On the other hand, including to many parameters can make the recursive calculation of derivatives (or "back-propagation") difficult.

The fully connected neural network used for this research has three hidden layers, since (Gu et al., 2020) conclude that a neural network performs the best if it contains three hidden layers. When another hidden layer is added, the performance of the network decreases. The input layer consists of 16 units, given that we have 4 variables and we are including 4 lags per variable. The three hidden layers have 16, 8 and 4 neurons, respectively. We choose the amount of neurons in each layer according to the geometric pyramid rule, which is introduced by (Masters, 1993). Each neuron in our neural network has the same activation function: the Sigmoid function. This function is used the most when it comes to machine learning. The function is defined as:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{7}$$

Let K^l be the number of neurons in each hidden layer, with $l = 1, 2.....L$ hidden layers. In our case $L = 3$. The recursive formula for the output at each neuron of the neural network is given by:

$$s_k^{(l)} = Sigmoid(s^{(l-1)'} \beta_k^{(l-1)}) \tag{8}$$

The final output is then:

$$y(x;\beta) = s^{(L-1)'} \beta_k^{(L-1)} \tag{9}$$

The network we created is complex and loaded with many parameters. Overall, the neural network has 560 parameters. Since this method is very nonlinear and not convex, the optimization becomes almost infeasible. A general solution is to train the network by making use of the stochastic gradient descent (SGD) method. This method uses an arbitrary subset of the data to estimate the gradient at every iteration of the optimization. This is an advantage over the standard descent methods, since those methods use the whole training sample. Consequently, this gives a huge boost to the optimization routine. At the same time the accuracy of the approximations decreases.

Another problem caused by the flexibility and parametrization of the network, is overfitting. An overfitted model is a model that contains too many parameters and is able to notice even the slightest variations in the training data, which worsens the forecasting performance. To prevent overfitting, regularization tools are used. For our model, we use four tools.
The learning rate is one of these tools. This method controls the step size of the descent and makes sure the calculation noise does not affect the descent.
The second regularization tool, is the ridge penalization. This is a general form of regularization, which reduces the multicollinearity and complexity of the model.
Another general regularization tool, is early stoppage. This method decreases the forecast errors in the data by updating the parameters at every iteration. The optimization is put to an end, when the forecast errors of the validation sample begin to increase.
The last regularization tool we use, is "batch normalization" (Ioffe & Szegedy, 2015), which normalizes the input for every hidden layer and smooths the objective function. This technique improves the speed and stability of the network. It minimizes internal covariate shift, which occurs when inputs of the hidden layers follow different distributions for the training and validation sample. With batch normalization, the network is able to use a higher learning rate, without vanishing or exploding gradients.

Figure 2 shows a representation of the neural network. The input variables are named $X_1 ... X_{16}$. The neurons are noted as S_n^l, where n is the number of the neuron and l is the number of the hidden layer.

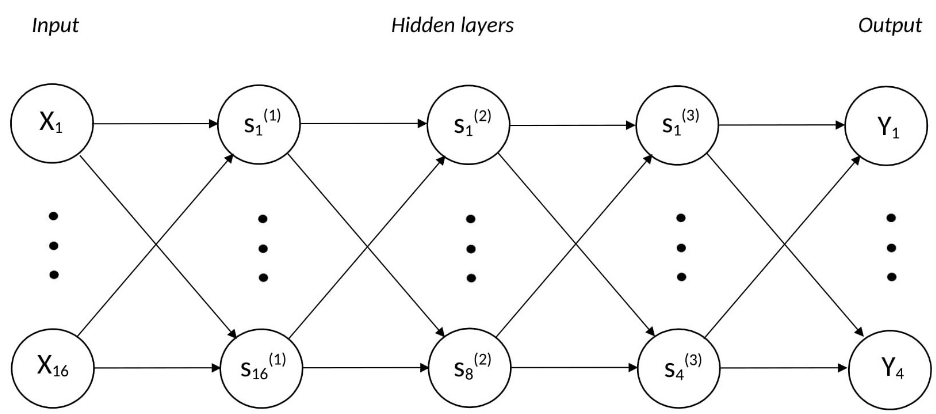

Figure 2: A representation of the neural network

For the coding of the neural network in Python, we used the packages (Abadi et al., 2015) and (Chollet et al., 2015).

4.2 Impulse Response

To examine the quantile-specific effects of variations in fiscal policy, we compute the impulse responses for the variables government spending and output. These impulse responses show us, how the variables respond to a 1 percent positive shock in government spending. We use the following procedure. For the quantiles $q = 0.1, q = 0.5$ and $q = 0.9$, we estimate the VAR models. After that we make use of a Cholesky decomposition to orthogonalize the covariance matrix of the errorterms. The order of the variables is extremely relevant here. We stick to the same order as we used for the regressions: g_t, y_t, τ_t, r_t. Finally we create a response vector by means of a loop which generates the responses at each point in time. Plotting this vector gives us the desired impulse responses. To compute this in MATLAB, we followed the steps and used the formulas described by Dr. Lutz Kilian in his book *"Structural Vector Autoregressive Analysis"* (Kilian, 2017).

4.3 Forecasting Performance

To analyse the predictive performance of both models, we perform rolling regressions. One of the measures used in the paper (Gu et al., 2020) to compare the methods, is the out-of-sample R^2. This R^2 is equal to one minus the summed squared error of the residuals (SSR) divided by the summed squared observations (SST). The formula for the out-of-sample R^2 of the quantile regressions is shown in the equation below. T_3 denotes the testing sample.

$$R^2_{oos} = 1 - \frac{\sum_{(i,t)\in T_3}(Q_q(z_t) - c(q) - \sum_{i=1}^{4} B_i(q) z_{t-i})^2}{\sum_{(i,t)\in T_3} z_t^2} \tag{10}$$

Another predictive performance measure is the mean squared forecast error (MSFE). This MSFE is the expectation of the summed squared error of the residuals. In this context, that is:

$$MSFE = \sqrt{\frac{\sum_{(i,t)\in T_3}(Q_q(z_t) - c(q) - \sum_{i=1}^{4} B_i(q) z_{t-i})^2}{N_{T_3}}} \tag{11}$$

Here N_{T_3} denotes the amount of observations in the test sample, which is equal to 44 in our case.

5 Results

In this section the estimates of the two models, the impulse responses for the QVAR model and the results for the out-of-sample forecasts are presented. In section 5.1, the results of the quarterly quantile regressions and the neural network are shown. The impulse responses are displayed in section 5.2. Section 5.3 contains the out-of-sample results using an expanding window.

5.1 Model Estimates

The model estimates consist of two parts. First, the fitted values for the quantile vector autoregressive model are shown. After that, the R^2 of the two models are discussed.

Fitted Values

Before we start with the comparison between the two models, we take a look at the fitted values. We perform quantile vector autoregressive (QVAR) regressions for the quantiles $q = 0.1$ (lower decile) and $q = 0.9$ (upper decile) and estimate the corresponding fitted values. To get an indication of their deviations from trend, we plot these fitted values together with the actual output. Since the actual output is denoted by y_t, the lower and upper decile are given by $Q_{0.1}(\hat{y_t}|z_{t-1}, ..., z_{t-4})$ and $Q_{0.9}(\hat{y_t}|z_{t-1}, ..., z_{t-4})$, respectively. Here $\hat{y_t}$ represents the fitted values for output. The graphs is shown in Figure 3, where the actual output is shown by the solid green line. The orange and blue line show the 0.1 and 0.9 quantiles of the conditional distribution of output, respectively.

From the figure we observe the parts in our data where the fitted values for the upper and lower decile are close to the actual output. These parts represent periods when a change in fiscal policy occurred, which had a large influence on output, since output was far from its mean. We notice that this mostly happens when output is strongly positive or negative and thus not at its peak. So, output is in its lowest or highest 10 per cent, when the economy is turning to a recession or expansion, respectively. However, output is close to its mean when the economy resides in one of the two regimes.

When we compare Figure 3 with Figure 1 from (Linnemann & Winkler, 2016), we see that the upper and lower decile are closer to the actual output in their paper. As mentioned before, we excluded one variable from the data set used by (Linnemann & Winkler, 2016), namely the ratio of government debt held by the public. This is because the "debt" variable worsens the results; the gap between the actual output and the fitted values for the upper and lower decile were larger than shown in Figure 3. This difference can be explained by the size of the data set, since there was no quarterly data available for the "debt" variable for the period 1955Q1-1965Q4. The appendix contains the graph of the fitted values where the "debt" variable is included in the QVAR model (Figure 5).

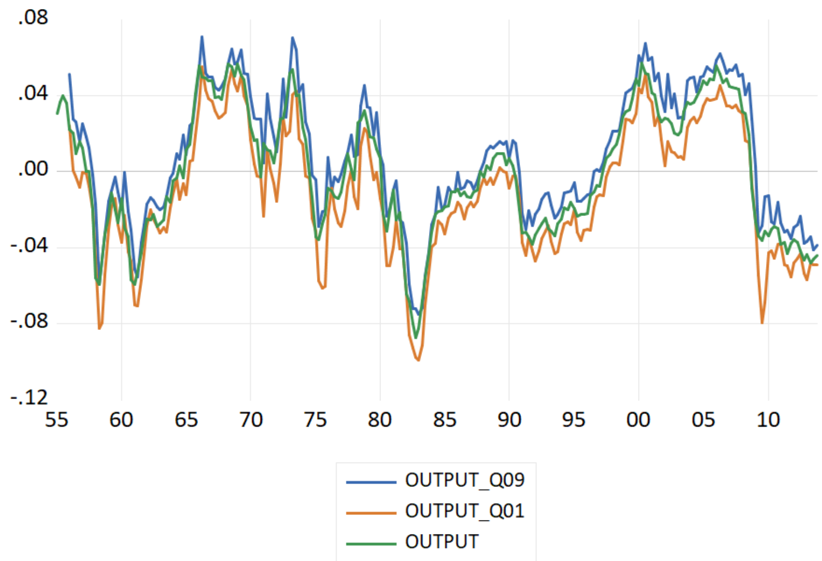

Figure 3: Actual output versus quantile forecasts

R squared

The table below shows the R^2 of the quantile regressions, performed at $q = 0.5$, and the neural networks. Since a VAR model can be estimated equation-by-equation, we performed 4 quantile regressions. Each regression consists of one dependent variable (mentioned left in the table), a constant and 16 explanatory variables (4 lags for each variable). To obtain these values for the neural networks, we constructed 4 different neural networks with one node in the output layer. An illustration of this network is included in the appendix (Figure 7). From the table we can conclude that the R^2 for the neural networks are higher than the R^2 for the quantile regressions. This means that the neural network explains a larger proportion of the dependent variable.

Table 2: R^2 of the models

	Quantile Regressions	Neural Networks
Government Spending	0.813	0.900
Output	0.782	0.941
Net Taxes	0.736	0.916
Interest Rate	0.521	0.722

When we compute the system as a whole, we get a R^2 equal to 0.513. The R^2 for the neural network with 4 nodes in the output layer, is equal to 0.718. Again, we obtain a larger R^2 for the neural network, which indicates that the neural network outperforms the quantile regression in terms of estimation of the nonlinear relation between the variables.

The R^2 for the quantile regression, where output is the dependent variable and $q = 0.1$, is equal to 0.760. When $q = 0.9$, the R^2 equals 0.773. These values are lower than the R^2 of the neural network, with the variable output in the output layer; 0.941. The R^2 for the neural network is still higher. However, this comparison is not completely fair, since we are comparing the R^2 at different quantiles for different methods.

5.2 Impulse Responses

To investigate the effect of fiscal policy expansions, we take a look at the quantile-specific impulse responses. These impulse responses are based on the QVAR models.

The impact of a positive 1 percent shock in the median of government spending on the variables output and government spending, for different quantiles ($q = 0.1$, $q = 0.5$ and $q = 0.9$) of output, can be obtained by the orthogonalized impulse responses shown in Figure 4. The order of the variable in this model is: g_t (government spending), y_t (output), τ_t (net taxes), r_t (interest rate).

From the graphs displayed in the first part of the figure (part a), it is visible that output responds positively to a shock in government spending. When we compare the three graphs in the first row, we oserve a difference in the responses. When output is in its lowest 10 percent ($q = 0.1$), the positive shock in government spending causes a strong return with a peak that lasts about 7 to 8 quarters after the change. At the median ($q = 0.5$), output reponds relatively weaker and shorter compared to the response at $q = 0.1$. There is an increase in the economic activity which is only visible till approximately 2 to 3 quarters after the shock. At the highest decile, $q = 0.9$, there is even a sign of a negative response after approximately 4 quarters. However, after about 11 quarters, the economy rises and output becomes positive again. Output responds differently for the three different quantiles. Consequently, we obtain a nonlinear relation.

When we compare these three graphs with the first row of Figure 2 from (Linnemann & Winkler, 2016), we observe that the responses at $q = 0.1$ and $q = 0.5$ are approximately the same. However, at $q = 0.9$, there is a difference between the responses we obtained: in their graph, when output becomes negative, it stays negative, but the response we observe becomes positive again after some quarters. A reason for this difference may be the omitted "debt" variable, which was mentioned earlier. However, the response we obtained stays in the 90 percent confidence interval shown in Figure 2 from (Linnemann & Winkler, 2016).

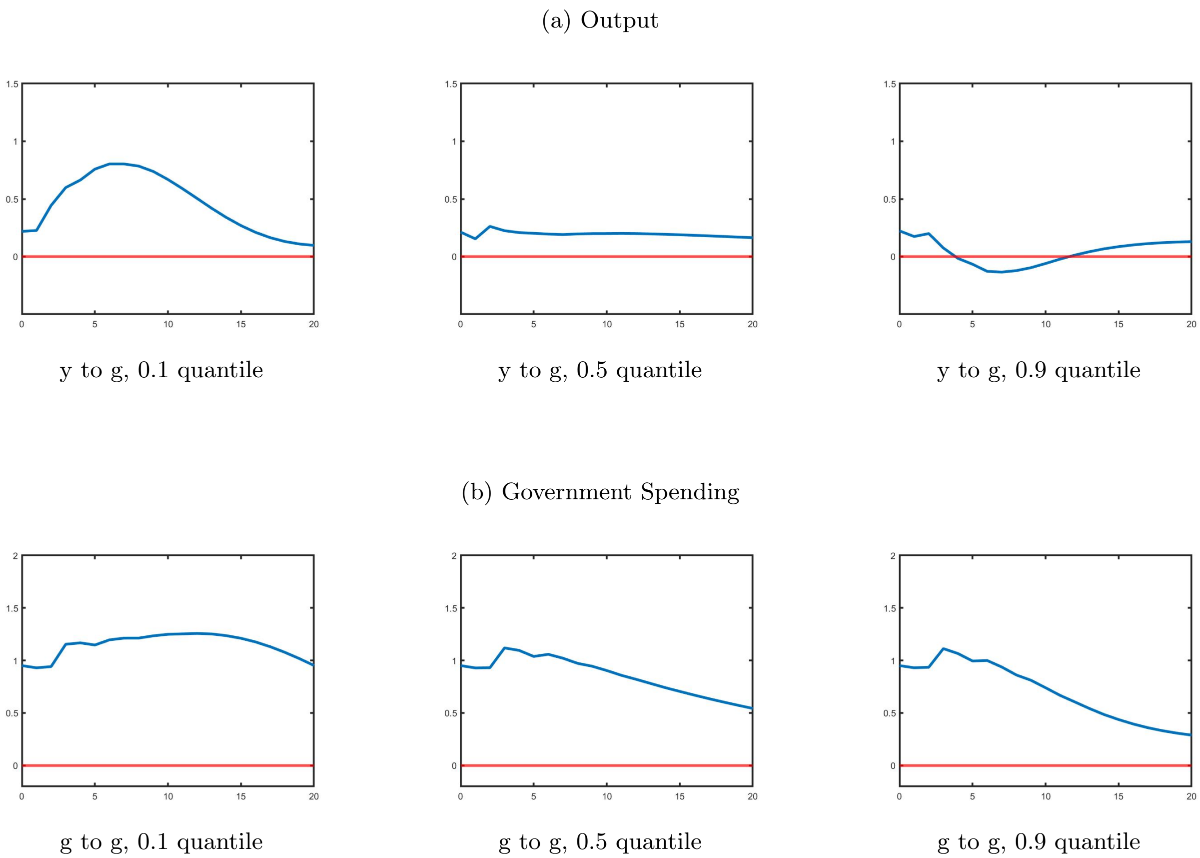

Figure 4 Impulse responses at different quantiles of the output distribution

Note: The solid lines show the responses of output and government spending to a 1 percent government spending shock, when the parameters are estimated at the 0.1, 0.5, and 0.9 output quantiles.

The second part of the figure (part b) displays how government spending responds to the fiscal expansion. When we compare the responses at different quantiles, we observe that government spending has a strong response when output is relatively low ($q = 0.1$). The responses of government spending for the 0.5 and 0.9 quantile of output are almost the same. Again, we observe a brief and relatively weak response at these two quantiles: after approximately 6 quarters there is no further increase in government spending.

Comparing these three graphs with the second row of Figure 2 from (Linnemann & Winkler, 2016), shows that the responses at the different quantiles are more or less the same. However, we notice a difference in the response at $q = 0.1$: in their graph the increase in government spending declines, but it still keeps increasing. However, from the corresponding graph shown in this paper, we can see that government spending decreases after about 14 quarters. Again, the omitted variable may be the reason behind this difference.

5.3 Rolling Regressions

Since both methods are performed on VAR models, we forecast the system equation-by-equation. We use an expanding window to predict the data one step ahead for 2003Q1-2013Q4, given the data from 1955Q1 till 2002Q4. The results are shown below.

Table 3 displays the out-of-sample R^2 for both models. We observe larger values for the neural networks than for the quantile regressions. This indicates that the neural network explains a larger proportion of the forecasted dependent variable. Also, the R^2 for Interest Rate is lower than the other values.

Table 3: Out-of-sample R^2 of the models

	Quantile Regressions	Neural Networks
Government Spending	0.834	0.920
Output	0.791	0.967
Net Taxes	0.724	0.890
Interest Rate	0.539	0.773

The out-of-sample R^2 for the QVAR model as a whole is equal to 0.527. When we forecast the neural network with 4 nodes in the output layer, we get an out-of-sample R^2 equal to 0.764. Again, this value is higher than the out-of-sample R^2 for the QVAR model.

Table 4 shows the mean squared forecast errors (MSFE) of the two models. We observe smaller values for the MSFE of the neural networks than the quantile regressions. This implies that the forecasts obtained by means of a neural network are closer to the actual data than the forecasts obtained by means of a quantile regression.

Table 4: MSFE of the models

	Quantile Regressions	Neural Networks
Government Spending	0.328	0.235
Output	0.397	0.214
Net Taxes	0.623	0.529
Interest Rate	0.701	0.535

The MSFE for the the QVAR model as a whole is equal to 0.642. For the neural network with 4 variables in the output layer, the MSFE is 0.498. Again, the MSFE for the QVAR model is higher.

6 Conclusion

The research question we stated in the introduction, is: *"Does a neural network outperform a quantile regression in terms of estimation and prediction of the nonlinear relation between US government spending and US output?"* First we estimated this relation. Then, to compare the predictive performance of the two models, we performed rolling regressions.

To estimate the nonlinear relation between fiscal policy and macroeconomic activity, we started with reproducing the quantile vector autoregressive (QVAR) models. For our second model, we used a fully connected, feed-forward neural network. This network estimates the same VAR model. We obtained larger values for the R^2 of the neural networks, compared to the R^2 for the quantile regressions. This indicates that a neural network outperforms a quantile regression in estimation of the nonlinear relation.

To predict the relation between government spending and output, we had a look at the out-of-sample forecasts. We made use of an expanding window, where we used the observations from 1955Q1 till 2002Q4 to predict the data from 2003Q1 till 2013Q4. The results show that the neural network explains a larger proportion of the forecasted dependent variable, since the out-of-sample R^2 are higher than the out-of-sample R^2 for the quantile regressions. Also, the MSFE for the neural networks are lower compared to the corresponding values for the quantile regressions, indicating that the neural network gives forecasts which are closer to the actual data.

Thus, we conclude that neural networks estimate and predict the nonlinear relation between US government spending and US output better than quantile regressions.

7 Discussion

This paper gives the impression that the machine learning method outperforms quantile regressions in terms of estimation and prediction of the nonlinear relation between fiscal policy and macroeconomic activity. However, this does not mean that machine learning methods do not have limitations.

For our research we used the nonlinear machine learning method neural networks. This method is complex and not transparent. The network requires a large data set in order to catch the nonlinear trend between the variables. The sample we used to train the modek contained 188 data points, which is very small. The neural network may have failed to capture the nonlinear behaviour in our data.

Another important point is the fact that we did not use the ratio of debt held by the public. When replicating the results of (Linnemann & Winkler, 2016), this causes differences. However, this does not affect the comparison between the two models used in this paper, since the same variables are used for both models.

We estimated the nonlinear relation between US government spending and US output at different quantiles of output by means of a QVAR model ($q = 0.1$, $q = 0.5$ and $q = 0.9$). However, for the neural network, we only estimated the relation at $q = 0.5$. For further research it might be interesting to include different quantiles of output for the neural network as well.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems.* Retrieved from `https://www.tensorflow.org/` (Software available from tensorflow.org)

Akimov, D. I. L., Azagouag. (2020). Machine learning catalysts in consumption growth tracking portfolios. *Erasmus School of Economics*.

Auerbach, A., & Gorodnichencko, Y. (2012). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy*, *4*(2), 1–27.

Blanchard, O., & Leigh, D. (2013). Growth forecast errors and fiscal multipliers. *American Economic Review*, *103*(3), 117–20.

Blanchard, O., & Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics*, *117*(4), 1329–1368.

Blanchard, O., & Perotti, R. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, *95*(1), 161–182.

Cecchetti, S. G., & Li, H. X. (2008). Measuring the impact of asset price booms using quantile vector autoregressions. *Brandeid University, Waltham*.

Chollet, F., et al. (2015). *Keras.* `https://keras.io`.

Feng, G., Giglio, S., & Xiu, D. (2017). Taming the factor zoo. *Chicago Booth research paper*(17-04).

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223-2273. Retrieved from `https://doi.org/10.1093/rfs/hhaa009` (2223) doi: 10.1093/rfs/hhaa009

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359 - 366. Retrieved from `http://www.sciencedirect.com/science/article/pii/0893608089900208` doi: https://doi.org/10.1016/0893-6080(89)90020-8

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.

Kelly, B., & Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, *186*(2), 294–316.

Kilian, L. (2017). *Structural vector autoregressive analysis.* Cambridge University Press.

Koenker, R., & Bassett, G. (1987). Regression quantiles. *Econometrica*, *46*(1), 33–50.

Linnemann, L., & Winkler, R. (2016). Estimating nonlinear effect of fiscal policy using quantile regression methods. *Oxford Economic Papers*, *68*(4), 1120–1145.

Masters, T. (1993). *Practical neural network recipes in c++.* Academic Press.

8 Appendix

Additional Figures

Figure 5 shows the fitted values for the lower and upper conditional output decile and the actual output. Here, we included the "debt" variable. The solid green line refers to actual output, the orange line refers to the lower output decile, and the blue line refers to the upper output decile. After comparing this figure to figure 3, we can conclude that the deviations are larger when the debt ratio variable is included.

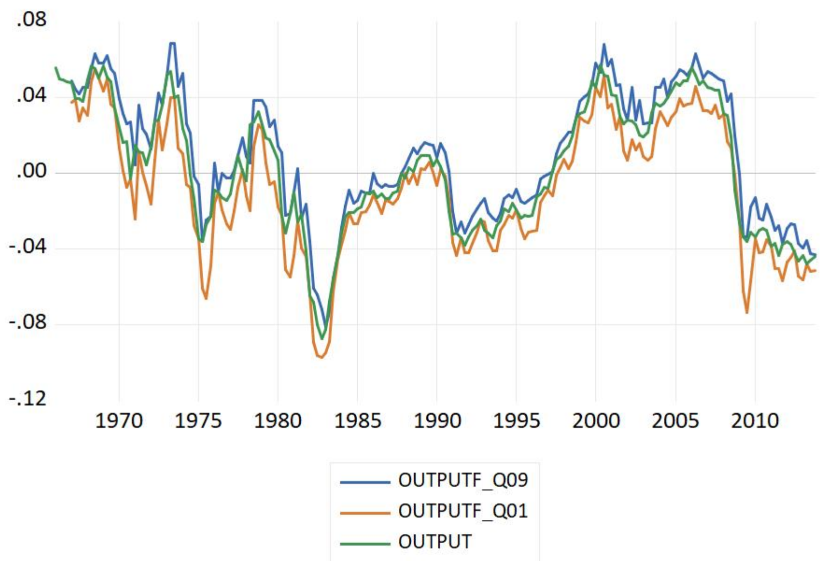

Figure 5: Actual output versus quantile forecasts, including the debt variable

Figure 6 gives an illustration of a neural network without any hidden layers. This network is the same as OLS.

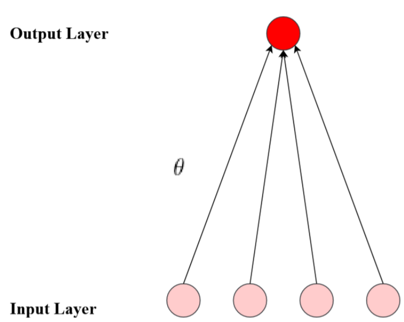

Figure 6: Neural Network without hidden layers, illustration from (Gu et al., 2020)

In Figure 7 the neural network, used to estimate the nonlinear relation equation-by-equation, is shown. This network has only one node in the output layer.

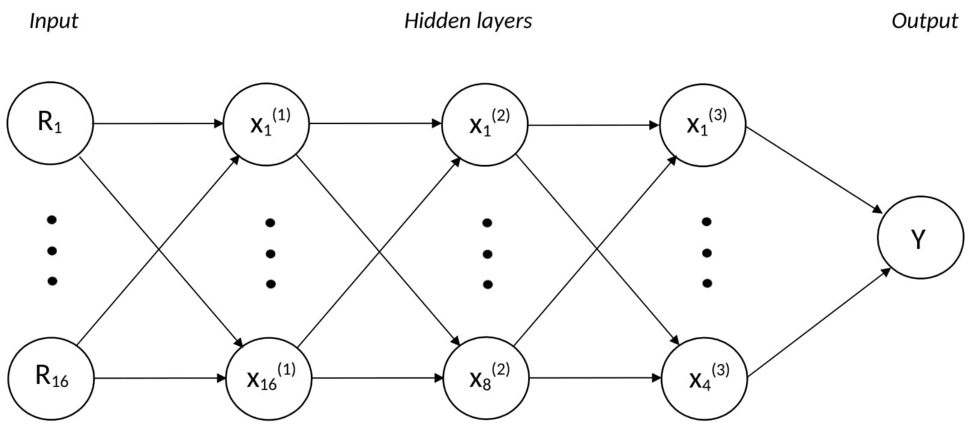

Figure 7: Neural Network with one variable in the output layer

Code

To obtain the results for this research, three softwares were used: EViews, MATLAB and PyCharm (Python). The attached ZIP file, contains the three folders with the code. Here, the folder Python contains three files, namely "NN", "QuantileRegressions" and "Expanding window". The last mentioned file is used in combinatio with the two models, in order to evaluate the predictive performance. Since we also used neural networks for the paper (Akimov, Azagouag, Djibuti, Ilyas & Lingsveld, 2002), the code for this research does contain parts of the seminar paper.