ALINA TÄNZER

The Effectiveness of Central Bank Purchases of long-term Treasury Securities: A Neural Network Approach

Institute for Monetary and Financial Stability
GOETHE UNIVERSITY FRANKFURT

WORKING PAPER SERIES NO. 204 (2024)

Institute for Monetary and Financial Stability
Goethe University Frankfurt
House of Finance
Theodor-W.-Adorno-Platz 3
D-60629 Frankfurt am Main
www.imfs-frankfurt.de | info@imfs-frankfurt.de

The Effectiveness of Central Bank Purchases of long-term Treasury Securities: A Neural Network Approach

Alina Tänzer[a,]

[a]*Institute for Monetary and Financial Stability, Goethe University Frankfurt, Theodor-W.-Adorno Platz 3, 60323 Frankfurt, Germany*

Abstract

Central bank intervention in the form of quantitative easing (QE) during times of low interest rates is a controversial topic. This paper introduces a novel approach to study the effectiveness of such unconventional measures. Using U.S. data on six key financial and macroeconomic variables between 1990 and 2015, the economy is estimated by artificial neural networks. Historical counterfactual analyses show that real effects are less pronounced than yield effects. Disentangling the effects of the individual asset purchase programs, impulse response functions provide evidence for QE being less effective the more the crisis is overcome. The peak effects of all QE interventions during the Financial Crisis only amounts to 1.3 pp for GDP growth and 0.6 pp for inflation respectively. Hence, the time as well as the volume of the interventions should be deliberated.

Keywords: Artificial Intelligence; Machine Learning; Neural Networks; Forecasting and Simulation: Models and Applications; Financial Markets and the Macroeconomy; Monetary Policy; Central Banks and Their Policies (JEL: C45, E47, E44, E52, E58)

1. Introduction

The effectiveness of unconventional policy measures conducted by the Federal Reserve (FED) during the great Financial Crisis is controversially discussed. Especially against the background of further balance sheet increases which recently happened due to the Corona crisis, quantitative easing (QE) is a hot topic. In this regard, the paper at hand provides new evidence on the effect of such measures, using information from Federal Open Market Committee (FOMC) announcements to construct a policy measure, and employing a novel approach using artificial neural networks (ANNs) as model setup.

Recent attempts to shed light on this topic made use of theoretical or empirical models to extract the reaction of the economy to balance sheet adjustments (see for example

*Corresponding author *Email address*: alina.taenzer@web.de

Gertler and Karadi (2013) and Kim et al. (2020)). While conventional theoretical models a-priori impose certain transmission channels of the policy, the economic interrelations estimated in vector autoregressive (VAR)-based analyses are bounded by their linear framework. However, as this policy tool is not frequently employed, and initially introduced during disruptive times, the underlying mechanisms may well encompass nonlinearities of unknown form. Therefore, this paper makes use of neural networks, which comprise a flexible and powerful tool which is purely data-driven. Except for the selection of input variables - which is virtually unlimited - the model does not rely on economic theory and hence no parametric cross-requirements need to be imposed. Building on the universal function approximation property (see Hornik et al. (1989)), the ANN represents any functional form to an arbitrary degree and thereby captures underlying coherences better than conventional models. Hence, an ANN-based setup may be especially well suited to discover the effect that balance sheet changes have on the financial sector and the real economy. To the best of the author‘s knowledge, this is a novel approach which has not been applied for policy evaluation so far.

In this paper, new evidence on the effect of Federal Reserve's large-scale Treasury securities (TS) purchases on the financial markets and the macroeconomy is provided. Based on novel techniques of artificial intelligence, a system of ANNs is estimated using six key variables: federal funds rate (FFR), 10-year yield, excess bond premium (EBP), inflation, GDP growth and a large-scale asset purchase (LSAP) measure[1]. Using information from FOMC announcements, an unconventional policy measure is constructed representing the one-year ahead level of TS on the FED's balance sheet. Estimating the ANN-economy for the period 1990:2015, and extracting the structural shocks thereof, a historical counterfactual analysis following Primiceri (2005) is conducted. Furthermore, impulse response functions for individual policies (the first, second and third round of QE during the Financial Crisis, colloquially denoted by QE1, QE2 and QE3) disentangle the effectiveness at different crisis depths. In order to assess the potential of using balance sheet adjustments as an alternative to conventional interest rate policy during normal times, counterfactual scenarios for pre- and post-crisis times are demonstrated.

As the results outlay, real effects are less pronounced than the effects on the financial sector, and even become negative towards the end of the analyzed period. This is obviously caused by positive short-run and negative long-run effects. Further, analyzing each policy intervention individually, the results indicate heterogeneity in the transmission of the policy. The economic stance prior to the QE measure influences its impact, which is larger the deeper the crisis. Obviously, the ANN reveals disproportionalities, which underline the benefit of relying on this modeling-method. In an economic sense, the findings indicate that QE should rather be conducted when the economy is in a deep recession. Besides the timing of the intervention, which is shown to be cru-

[1]Please note that in the following, when talking about LSAP, it is referred to the central banks' purchases of long-term Treasury Securities, which is only one part of the total LSAP measures conducted by the central bank.

cial for its power, its volume should be subject to discussion, as GDP growth (inflation) is increased by around 1.3 pp (0.6 pp) only, adding up the effects of all interventions on impact. Yielding rather small real effects, the findings range at the lower bound of research in this field.

This paper first of all relates to the field of literature addressing the same question of QE effectiveness. There are on the one hand theoretical DSGE models which inherit different transmission mechanisms of the policy (see e.g. Gertler and Karadi (2013), Chen et al. (2012), Chung et al. (2012), Carlstrom et al. (2017), Ellison and Tischbirek (2014)). On the other hand, several empirical models measure the effect of QE on the yield curve or the real sector (see Baumeister and Benati (2010), Gagnon (2010), Neely et al. (2010), Swanson (2011), Hamilton and Wu (2012)). In a previous paper, we provide an overview of the discovered effects revealed in the papers mentioned, and further challenge multiple DSGE models under a harmonized policy tool (Tänzer and Wieland, 2021). The heterogeneity of results found in the literature provokes the question, whether these approaches are suitable to answer the question at hand. Moreover, this paper connects to literature applying with neural networks in a macroeconomic forecasting setup. To name only some of them, Smalter Hall and Cook (2017) for example compare the forecast performance of several deep neural networks to that of the professional forecasters' survey (SPF). Verstyuk (2020) uses networks with memory of various sizes to predict US Data on five key macroeconomic variables. In a previous paper (Tänzer, 2021), which also includes a more detailed overview on the neural network forecasting literature, the linkage between conventional and ANN-based approaches is created by comparing their forecasting performance in a multivariate macroeconomic setup. The robustness of ANN-based predictions over crisis periods is shown, as well as its superiority over DSGE and Bayesian VAR, especially in recent times. Further, there exists literature on nonlinear relationships within the economy, which consider convex Phillips or IS curves (Dolado et al. (2004), Dolado et al. (2005), Coibion and Gorodnichenko (2015)). These findings further justify the use of ANN as economy representation. In terms of designing the policy measure and with regard to the setup, this paper is most closely related to (Kim et al., 2020), who answer the same question in a VAR analysis and also conduct a historical counterfactual as well as several impulse response functions.

2. Methodology and Data

2.1. *Artificial Neural Networks*

In this study, a novel approach to data-driven effectiveness analyses is introduced, which is based on neural networks. While the idea of ANNs dates back to the 1940's, they experienced an increasing amount of attention linked to the processing power of computer technology and data availability in the early 2000's. Continuous innovations improved training efficiency, reduced the risk of overfitting and rendered model training a feasible task. This led to widespread applications of neural networks in many fields.

2.1.1. Basic Concept

Neural networks can be imagined as directed graphical models, in which information flows from inputs via a specific structure to a target output. Within the *structure*, nodes represent operations on the data as it flows from input to output. For economists, who are used to focus on specifying the model equations and the relevance of data inputs, the approach to neural networks will sound different. Here, model architectures are discussed, which refers to the configuration of the network structure, i.e. the number of nodes, the number of so called layers, the interconnections between these nodes and layers, and the nature of operations performed at each node.

The basic architecture of a neural network model consists of three distinct sets of nodes collected each in one layer. First, a set of nodes representing model inputs (Input Layer), second a set of computational nodes (Hidden Layer) and third a set of nodes constituting model outputs (Output Layer). In deep learning models, the number of Hidden Layers and nodes per layer is increased to an arbitrary size. A basic neural network structure with only one hidden layer is schematically displayed in Figure 1. The Input Layer contains all variables of interest as well as their lags. This data is weighted by individual parameters collected in vector ω_i and fed to J neurons (also called nodes), stacked together in the Hidden Layer. These nodes perform transformations on the weighted sum of inputs, complemened by a bias α_j, according to the transfer function $G(\cdot)$. The resulting values are further processed, again weighted with parameters collected in vector ω_j and assembled - adding another bias α_0 - to produce the final outputs. With this structure, the network is essentially similar to a Generalized Linear Model (GLM). During network estimation, input data as well as target output data is provided to the algorithm, and the weights and biases are determined such that a loss criterion (e.g. the mean squared error) is minimized. This training process is conceptually different to conventional econometric approaches which employ for example maximum likelihood estimation. Instead, the researcher can choose from various non-parametric training algorithms which rely on back-propagation and gradient descent for example.

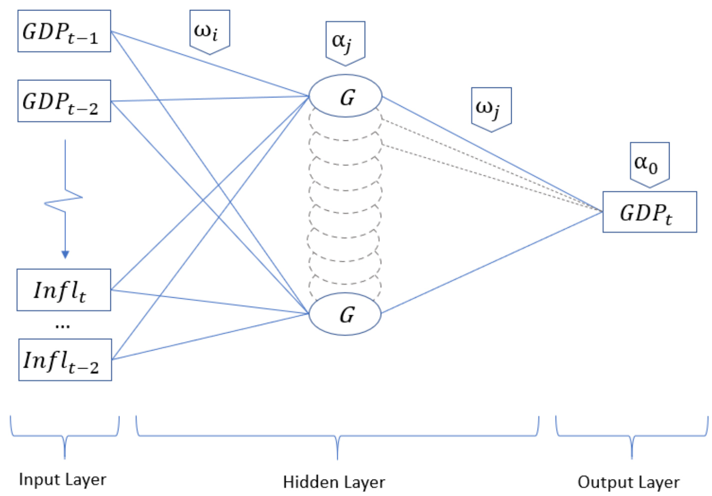
The ANN can also be respresented by the following equation, which relates the variable of interest m_t to the function $\hat{f}^m$:

$$m_t = \hat{f}^m = \alpha_0 + \sum_{j=1}^{J} \omega_j G\left(\omega_i' z_t^m + \alpha_j\right). \tag{1}$$

Here, z_t is the input vector with explanatory input variables which dependson the output variable m. The weight-parameters to be estimated are collected in vectors ω_j, with $j = 1, ..., J$ representing the chosen number of neurons in the Hidden Layer, and ω_i, with $i = 1, ..., I$ indicating the number of input variables, i.e. the length of z; also the respective “biases" α_0 and α_j are estimated.

One major advantage of neural networks, is that throughout the training process, the

Figure 1: ANN Scheme



model identifies which information (nodes) is relevant for the prediction[2], which provides the researcher with the freedom to be less selective regarding the data supplied into the model. Another crucial characteristic of neural networks is their universal approximator property (Hornik et al., 1989). It indicates, that any unknown function can be approximated arbitrarily well by a linear combination of transfer functions $G(\cdot)$, such that $|H(z_t) - \sum_{j=1}^{J} \omega_j G(\omega_i' z_t)| < \delta$ with J being finite and $\delta \in \mathbb{R}_{>0}$. This constitutes an additional advantage over conventional models, as the researcher does not have to choose a functional form for the interrelation since this specification is data-driven. Besides these powerful advantages, employing neural networks has the drawback of being locally identified only, and parameters lack economic interpretability. Consequently, their field of usage is mainly in conducting forecasts. Nevertheless, models which generate precise predictions may also be used for applications for which generally other econometric methods like VARs are employed. Important for using ANNs in such a macroeconomic context, is that the evolution of certain variables as reaction to changes in other variables is more important than *how* this reaction is generated. While taking partial derivatives can shed light on the underlying interrelations to some extent, one can also draw partial dependence plots, marginalizing over some variables. Furthermore, novel algorithms are able to calculate the partial variables' input importance for an output variable. Hence, these ideas can at least partly circumvent neural networks' *black-box* characteristic, which drives macroeconomists' scepticism against these novel tools.

To sum up, networks generally differ with respect to their network architecture in terms of its size (hidden layers and neurons per layer), the employed transfer func-

[2]With prediction, it is referred to the mapping of inputs to the target outputs, which is essentially equivalent to a prediction.

tion(s), the interconnection between nodes, the training procedure employed to estimate weights and biases, and many more settings which can be individually adjusted. Next, the employed model architecture is introduced.

2.1.2. Model Architecture

For this forecasting comparison a fully connected feed forward nonlinear autoregressive neural network with external inputs is chosen. The term *fully connected* means that there is a connection between every node on every layer. *Feed forward* characterizes the direction, in which data is transferred through the network, being from the Input through the Hidden towards the Output Layer only (no reverse movement). The term *autoregressive* means that lags of the output variable are also included as input data, here it is allowed for a lag length of 2, following Kim et al. (2020). The *external inputs* refer to input data which is not used as output. This setup is chosen as it is well suited for timeseries estimation and forecasting. Furthermore, in order to keep it simple and deviate as little as possible from conventional (VAR-based) methods, the employed network is designed with only one Hidden Layer. While more sophisticated and deeper networks might perform better in terms of forecasting, this simple network still fits the data sufficiently well and consequently constitutes a suitable benchmark model.

The analysis in this paper is characterized by the key variable (the balance sheet level of the central bank), which relevantly changes only during 2009 and 2015. Hence, the length of the dataset is somehow limited, although one can use data prior 2009 for estimating the remaining equations (see Section 2.2). Hence, one key issue when selecting the algorithm to be employed, is that no further reduction of the (training-) dataset is necessary. Typically, a network is trained using backpropagation, which relies on supervised learning, deploying a gradient descent method to reduce a chosen error function (e.g. mean suqared error). One caveat of this technique is the potential for overfitting, which leads to a loss of generalization due to fitting of the noise. Some algorithms control overfitting by constantly testing the generalization capabilities in a validation set. This obviously shrinks the dataset available for training. There are, however, alternatives to these approaches. In this regard, the Levenberg-Marquardt algorithm with Bayesian regularization (Foresee and Hagan, 1997) is employed to estimate the neural network's weights and biases. Bayesian regularization, developed by MacKay (1992) and transferred to neural networks by Foresee and Hagan (1997), is a technique to counteract overfitting without a validation set. While in general, during estimation, the mean squared errors E_D are reduced, this method aims at also shrinking the employed weights by expanding the objective to $F = \beta E_D + \alpha E_W$, where E_W is the sum of squares of the network weights and biases. The parameters α and β determine the relative importance of function approximation and generalization and are optimized through Bayesian methods. The required Gauss-Newton approximation of the Hessian matrix is achieved applying the Levenberg-Marquardt optimization algorithm. This procedure reduces the potential for arriving at local minima and thereby further incresases the generalizability of the network (Ticknor, 2013).

For the choice of the transfer function $G(\cdot)$, the author relies on a rectified linear form $ReLU(x) = max(0, x)$, which gained popularity in recent years (Glorot et al., 2011). Compared to sigmoid or similar activation functions, it allows faster and more effective training of neural networks using complex datasets. Next, the number of nodes J (also called the *width* of the network) has to be determined. In order to fulfill the universal approximator property in a setting with a single hidden layer and *ReLU* activation functions, a range for the minimum width is provided by Hanin and Sellke (2018)[3]. Given the number of inputs and output variables in this application, the minimum width accounts for 11 to 18 nodes. However, the required number of nodes still depends on the complexity of the underlying data. Bayesian regularization has the advantage of providing a measure of how many network parameters/nodes are effectively being used. Hence, in a pre-training, the required number of nodes is determined, training the network in a loop of up to 30 nodes. The number of nodes is said to be large enough, when the *number of effective parameters* is not increased further by adding more nodes to the structure. The resulting number of nodes is chosen as optimal network width and used in the actual training session. Here, the network is trained several times (in this case for 30 trials), because each trial starts with a different initialization of weights and biases which influences the network's performance. Out of the resulting 30 trained networks, the best one is selected according to the minimum mean squared error over the whole training data set.

2.2. Data and LSAP Measure

Data. The employed dataset covers monthly U.S. data from January 1990 to December 2015. While many authors, e.g. Gertler and Karadi (2015) use larger samples, starting in 1979, this paper follows Kim et al. (2020) in selecting a later starting point due to underlying structural changes during the years 1979 to 1990. There is for example evidence for a flattening of the Phillips curve during these years as shown e.g. by Blanchard et al. (2015). Hence, especially the inflation equation is more consistently estimated with the employed short sample.

The estimation of the ANN is based on six monthly variables - the Federal Funds Rate (FFR), the 10-Year Treasury yield, the inflation rate based on the consumer price index (CPI), the growth rate of industrial production (equivalent to GDP growth), the excess bond premium (EBP) based on Gilchrist and Zakrajšek (2012), which is equivalent to a credit spread, and the LSAP measure, which is described below. With this set of variables, this paper follows recent empirical research on the effectiveness of central bank balance sheet changes (Kim et al. (2020) and Gertler and Karadi (2015)). Please

[3]According to Hanin and Sellke (2018),"feed-forward neural nets with a single hidden layer can approximate essentially any function if the hidden layer is allowed to be arbitrarily wide." This result holds for a variety of activation functions, including ReLU. In detail, it states that any continuous function $f : [0,1]^{d_in} \to \mathbb{R}^{d_{out}}$ can be approximated arbitrarily close by a ReLU net N with input dimension d_{in}, output dimension d_{out} and minimum layer width ν: $|f(x) - f_N(x)| \leqslant \varepsilon$, with $\varepsilon > 0$. Calculating ν according to $d_{in} + 1 \leqslant \nu_{min}(d_{in}, d_{out}) \leqslant d_{in} + d_{out}$, gives the minimum width for the universal function approximation property to hold.

refer to the Appendix for more information on data sources and transformation.

LSAP Measure. In terms of unconventional monetary policy, the focus lies on Treasury Security (TS) purchases of large scale, as a representative of balance sheet increasing interventions by the FED. The LSAP measure is defined as the one-year ahead expected amount of TS on the Federal Reserve's balance sheet. More specifically, FOMC announcements are collected and the information about Treasury purchases is extracted. In this procedure, the author sticks to the approach in Tänzer and Wieland (2021). Furthermore, the current level of TS holdings by the central bank is known. With this information set, the one-year ahead expected TS-Balance sheet level is constructed by simply adding up both values. Certainly, purchases of other assets such as mortgage backed securities (MBS) are omitted here. Contrary, for example Kim et al. (2020) construct their LSAP measure based on all assets purchased by the central bank. This implies that the effect measured in this analysis only captures a fraction of the whole intervention and thus the fraction of the total effect related to purchases of TS. Furthermore, one could include survey data on central bank security holdings, which refines the expectations (this approach is chosen by Kim et al. (2020)). However, for simplicity it is abstracted from this.

2.3. *Estimation Strategy*

The estimation of the ANN is conducted in a recursive manner. This is equivalent to estimating a recursive VAR and hence, an ordering of the variables needs to be specified. The paper follows Kim et al. (2020) by ordering the equations as follows: FFR (i_t), LSAP (b_t), Yield (d_t), inflation (π_t), GDP growth (y_t) and EBP (p_t). Hence, the economy is represented by six individual networks, which all have the structure similar to Equation (1), with $m \in \{i, b, d, \pi, y, p\}$:

$$i_t = \hat{f}^i(i_{t-1}, i_{t-2}, d_{t-1}, d_{t-2}, \pi_{t-1}, \pi_{t-2}, y_{t-1}, y_{t-2}, p_{t-1}, p_{t-2}) + \varepsilon_{1t} \tag{2}$$

$$b_t = \hat{f}^b(i_t, i_{t-1}, i_{t-2}, b_{t-1}, b_{t-2}, d_{t-1}, d_{t-2}, \pi_{t-1}, \pi_{t-2}, y_{t-1}, y_{t-2}, p_{t-1}, p_{t-2}) + \varepsilon_{2t} \tag{3}$$

$$d_t = \hat{f}^d(i_t, i_{t-1}, i_{t-2}, b_t, b_{t-1}, b_{t-2}, d_{t-1}, d_{t-2}, \pi_{t-1}, \pi_{t-2}, y_{t-1}, y_{t-2}, p_{t-1}, p_{t-2}) + \varepsilon_{3t} \tag{4}$$

$$\pi_t = \hat{f}^\pi(i_t, i_{t-1}, i_{t-2}, b_t, b_{t-1}, b_{t-2}, d_t, d_{t-1}, d_{t-2}\pi_{t-1}, \pi_{t-2}, y_{t-1}, y_{t-2}, p_{t-1}, p_{t-2}) + \varepsilon_{4t} \tag{5}$$

$$y_t = \hat{f}^y(i_t, i_{t-1}, i_{t-2}, b_t, b_{t-1}, b_{t-2}, d_t, d_{t-1}, d_{t-2}, \pi_t, \pi_{t-1}, \pi_{t-2}, y_{t-1}, y_{t-2}, p_{t-1}, p_{t-2}) + \varepsilon_{5t} \tag{6}$$

$$p_t = \hat{f}^p(i_t, i_{t-1}, i_{t-2}, b_t, b_{t-1}, b_{t-2}, d_t, d_{t-1}, d_{t-2}, \pi_t, \pi_{t-1}, \pi_{t-2}, y_t, y_{t-1}, y_{t-2}, p_{t-1}, p_{t-2}) + \varepsilon_{6t}. \tag{7}$$

Due to the recursive formulation, the vector of input variables z_t^m differs for every variable in m. In order to ensure continuity, the LSAP measure is normalized to zero before 2008, i.e. the level of TS holdings prior the first intervention is subtracted from the series. One can hereby ensure that the increase in the balance sheet in 2009 captures only the expected increase due to the first round of QE. For the same reason, the ANN for

LSAP is estimated for the years 2009:2015 only (Equation (3)). The ANN representing the FFR is estimated over the whole sample (1990:2015) but does not contain the LSAP measure as input variable (Equation (2)). The thought behind this approach is that the central bank does not react to its own unconventional policy when setting the interest rate, but rather to the induced economic stance. Considering the whole sample in the nonlinear network further has the advantage of endogenously capturing the zero lower bound during the Financial Crisis. All other ANNs take advantage of the whole set of input variables over the whole sample period.

3. Results

3.1. Historical Counterfactual

In order to evaluate the effect of unconventional policy measures in the form of TS purchases by the FED, a historical counterfactual analysis is conducted. Taking into account the variables' dynamics, the economic behavior under the absence of QE can be contrasted against the actual evolution according to the data. Relying on the approach introduced by Primiceri (2005), the author uses the estimated ANN economy (Equations (2) to (7)) and the corresponding extracted estimated structural shocks (ε_{ht}, with $h = 3, .., 6$) to simulated the economy under different policies. Specifically, the central bank balance sheet is assumed to be kept at its pre-crisis level of zero, while the FFR is held constant at the ZLB. Since balance sheet measures are only used during the financial crisis, the counterfactual simulation period lasts from 2009 to 2015. The first two panels in Figure 2 show the exogenous paths of counterfactual monetary policy versus actual.

The implicit assumption during this analysis is that the market holds its beliefs about the systematic component of monetary policy unchanged[4]. Hence, it is subject to the Lucas (1976) critique as the market's beliefs about the FED's reaction function might have adjusted under a tighter policy. However, following Kim et al. (2020), this analysis may still be of interest, for example to measure the impact of prolonged obstacles to these balance sheet programs, in case they are not *perceived* to be permanent.

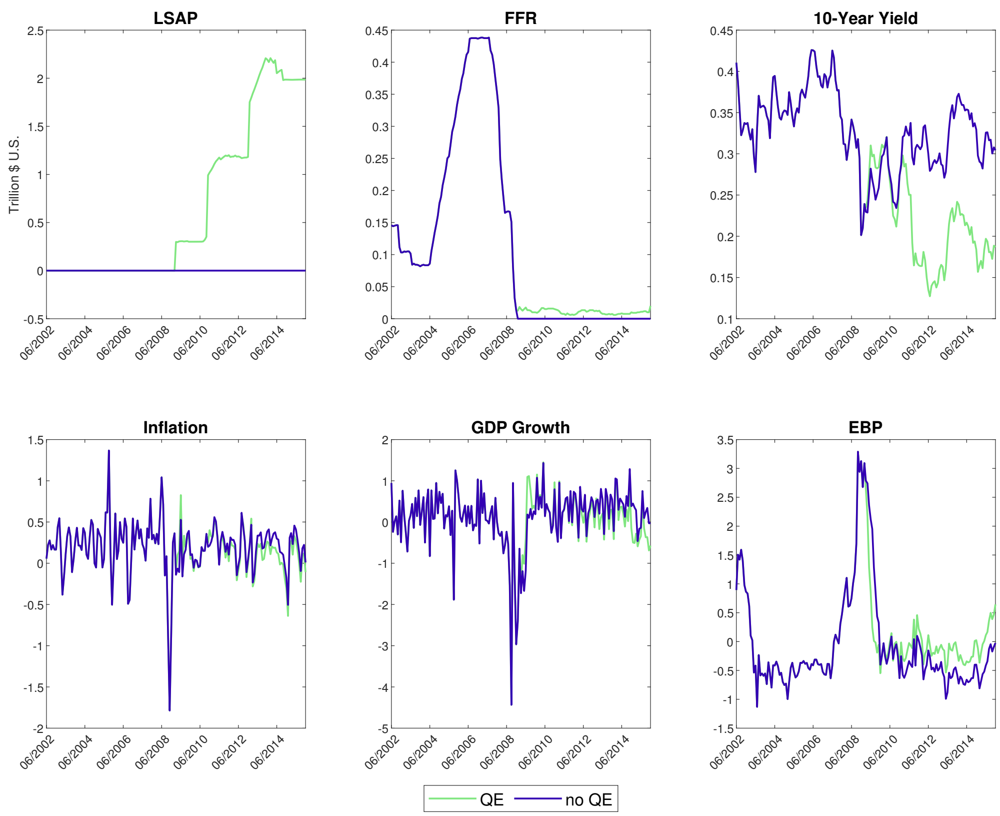
The remaining panels in Figure 2 provide endogenous counterfactual paths for the respective variables. While the 10-Year Yield including QE is initially slightly higher than without policy intervention, it is significantly reduced in subsequent periods. At its peak, QE reduces the Yield by 15 basis points. Contrary, the excess bond premium is initially reduced below non-QE levels, and afterwards a balance sheet increase generates EBP levels close to zero. The counterfactual levels of inflation and GDP growth show rather small differences towards actual, initially indicating higher levels under asset purchases and lower levels towards the end of the simulation period.

Figure 3 shows the differences Δ^{HCF}[5] between actual (QE intervention) and counter-

[4]Using the ANN economy, the estimated parameters are taken as given during this analysis and possible adjustments of the market's reaction to a change in monetary policy are thus not considered.

[5]HCF shall denote Historical CounterFactual analysis.

Figure 2: QE versus No QE Intervention



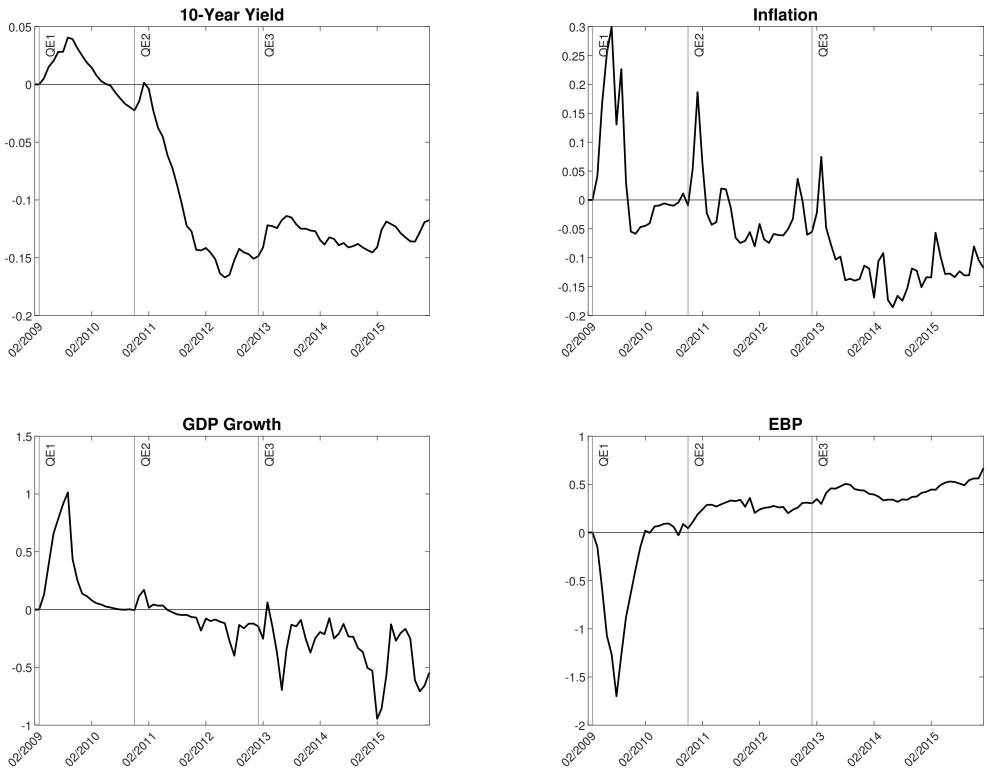
Note: This figure shows actual (including QE) and counterfactual paths (without QE) for all variables. All variables are in %.

factual (no QE intervention) during the relevant period 2009:2015, i.e.

$$\Delta^{HCF} = m_t^{HCF_{QE}} - m_t^{HCF_{noQE}}. \tag{8}$$

The first panel shows that following each QE announcement, the long-term yield is initially increased (by 4 bp after QE1, by about 3 bp after QE2 and by approximately 4 bp after QE3). In the longer term, the desired reduction in long-term yield is generated, being at its maximum 15 basis points below the level without asset purchases. As seen before, the EBP is reduced by 1.6 percentage points following QE1, however later QE measures only generate marginal reductions and the long-term trend shows rather higher values (about 0.4 pp) with than without QE. While this seems counterintuitive, the level of EBP is below zero during the observed period (see Figure 2). Hence, an increase in its level closes the negative excess premium after 2010. The real sector respresented by GDP growth is affected positively by each QE intervention, producing a 1 pp. increase after QE1, a 20 bp increase after QE2 and a 30 bp increase after QE3. Similarly, inflation is increased by 30 bp by QE1, while the effects shrink to 20 bp (QE2) and

Figure 3: Δ^{HCF} (QE versus no QE Intervention)



Note: This figure depicts the differences between actual and counterfactual paths. All variables are given in percentage points.

then 12 bp (QE3). The longer term levels of both variables, GDP growth and inflation, are lower assuming policy interventions (on average around -0.12 pp. for inflation and -0.5 pp. for GDP growth). Hence, one can see positive short-term effects of TS purchases, which decline with the interventions, as well as negative long-term consequences.

A similar analysis in a VAR-setup is conducted by Kim et al. (2020), measuring the counterfactual when no QE3 is implemented. They find the yield reduction to amount up to 1 pp, while the results at hand indicate a yield reduction of 1.8 pp for no balance sheet expansion at all[6]. Further, the EBP in the VAR counterfactual is about 1.5 pp lower without QE3, whereas the same effect appears only under all expansion programs. The effect on the real economy (GDP increased by 10 pp and inflation by 0.65 pp) is however much larger than those found in this paper.

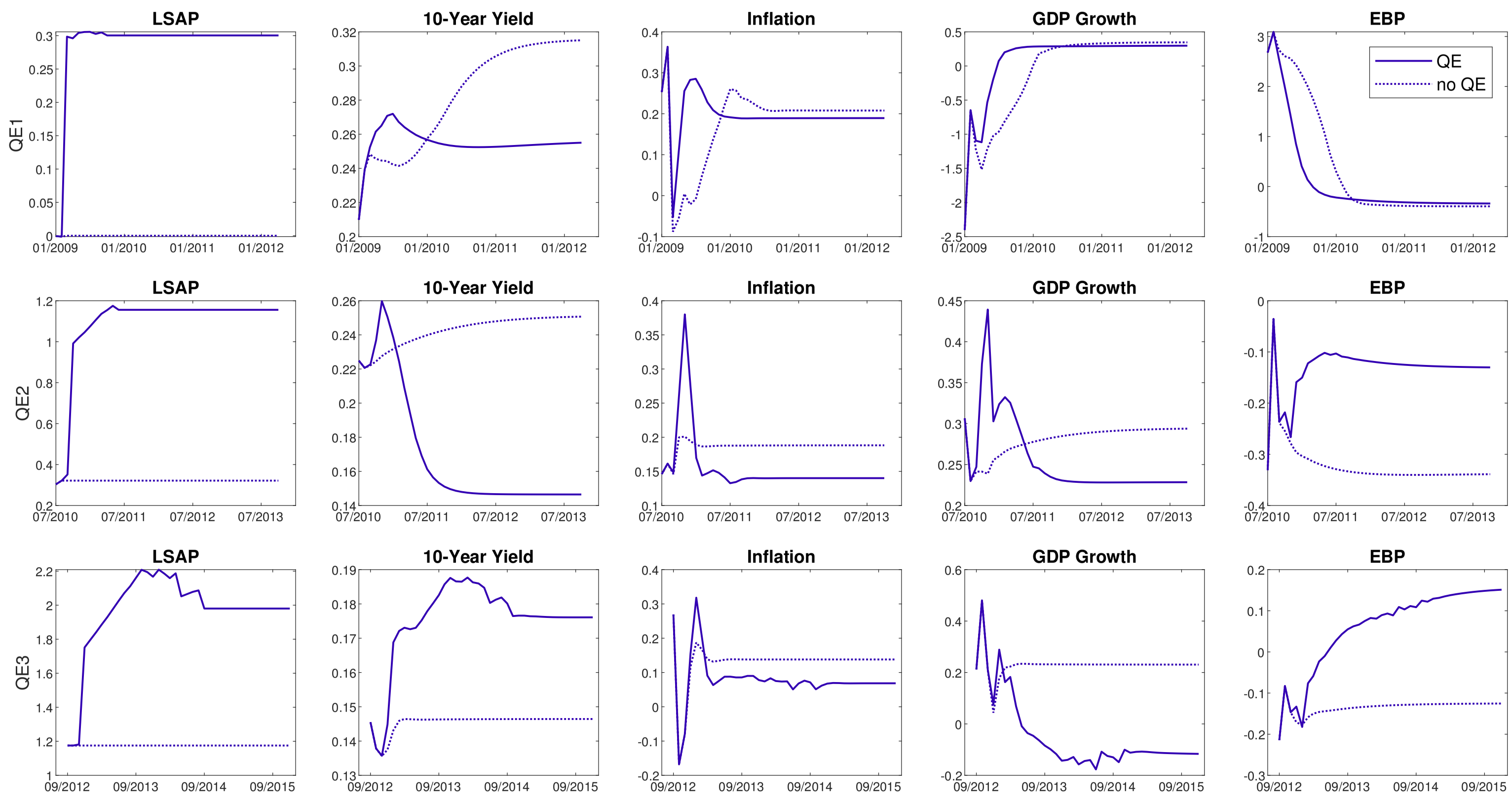
[6]Since the 10-year yield is used in a monthly specification, taking the difference between actual and counterfactual paths times twelve leads to the year on year yield which is shown in Kim et al. (2020)

3.2. Impulse Response Functions

For the purpose of disentangling the economic reaction to each unconventional policy intervention, impulse response functions are calculated. Specifically, the transition from the economic stance right before the intervention to stabilized levels is shown, with the balance sheet following the prescribed increasing or sustained paths . The FFR is again held constant at its ZLB. This analysis is different to the standard VAR-based or other model-based impulse response functions, which plot deviations from steady state levels. Nevertheless, the employed exercise provides relevant insights, since the impact of QE at different points in time, i.e. within differing crises depths, can be contrasted. The reason for differing effects of QE depending on the economic stance originates from the nonlinear structure of the economy.

Figure 4 provides results for this exercise, each row contains results for the respective QE intervention and each column contrasts a specific variable's path. As before, the two lines indicate the evolution with and without policy intervention. Considering QE1 first, the balance sheet is increased by US$ 300 bil., increasing the 10-year yield initially, but reducing it relative to no QE in the long-run. Hence, the desired yield effect occurs approximately one year after the intervention. This relates to the definition of the employed QE measure, which mirrors the expected 12 months ahead balance sheet level. Obviously, the effect on the yield curve occurs on impact of the intervention itself and not ex-ante driven by the policy announcement. Further, QE1 induces the EBP to decrease earlier. Looking at these two variables following QE2 the yield is also reduced in the long-term, whereas EBP shows higher levels under the balance sheet increase. As the premium is at negative values here, it is reduced through QE. In 2012, the policy becomes even more loose, with a third balance sheet increase to a level of US$ 2 tril. The 10-year yield is already at lower levels now (starting from 0.14% instead of 0.22% in 2009), and in this specific situation, QE raises the yield compared to the economic path without intervention. As for the second intervention, QE3 also raises the EBP compared to tighter monetary policy, in this case however above zero. Still, the long-term level of the EBP is closer to zero than the counterfactual.

Figure 4: Transition of Individual QE Intervention against no Policy



Note: This figure shows the transition from the economic stance right before each intervention under QE and no QE. All variables are in%.

Before talking about the key variables, lets take a look at Figure 5 to identify differences in terms of the magnitude of effects between interventions. This plot depicts the delta between paths under QE and no QE of Figure 4, i.e.

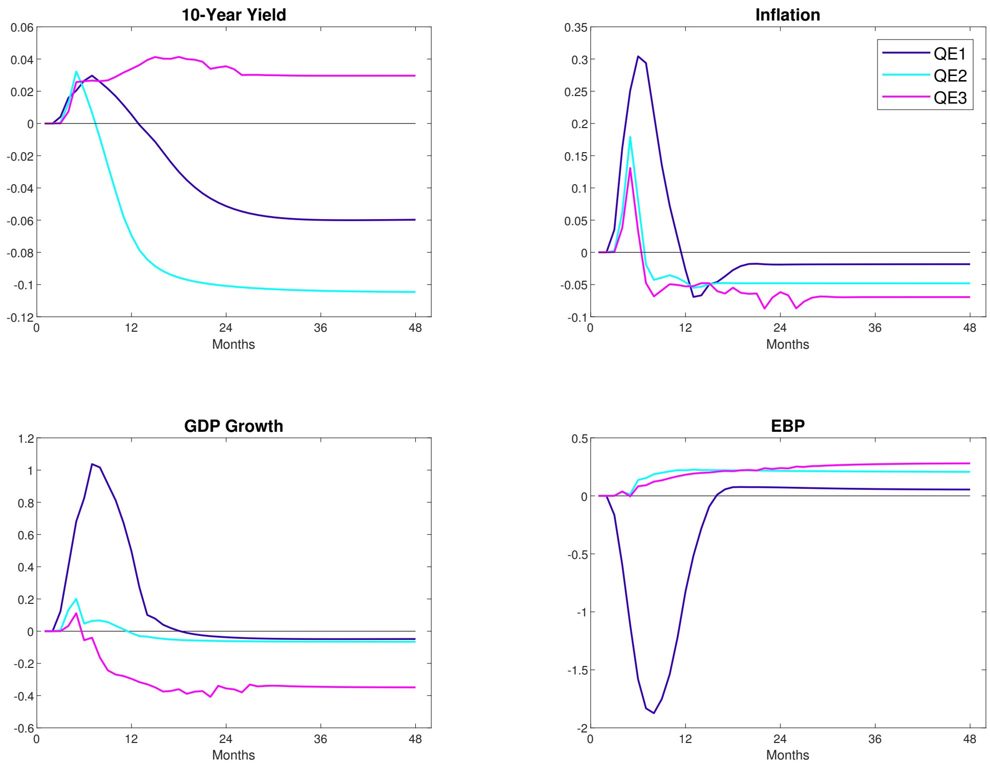
$$\Delta^{QE_x} = m_t^{QE_x} - m_t^{noQE_x} \tag{9}$$

with $x \in \{1,2,3\}$ and $m \in \{i, b, d, \pi, y, p\}$ as defined above. The 10-year yield is most effectively reduced by QE1 (around 6 bp in the long-run). While QE2 - which has about three times the size of QE1 - produces a reduction of about 10 bp, which does not relate to its larger volume. One can clearly deduce, that having recovered from the deepest points of the Financial Crisis, the balance sheet increase becomes less effective in terms of yield changes. In 2012, the yield is even increased (although only slightly by about 2 bp) when easing the policy further. This plot further shows, that the announcement of the balance sheet increase initially has the opposite effect than desired, as the yield is increased slightly. Following QE1 it takes about twelve months until the negative trend occurs, whereas after QE2, the desired negative path is established already after about 8 months. This might indicate that yield effects rather materialize when the balance sheet is changed and not following the announcement only. Furthermore, the yield decrease happens rather gradually and, at least for QE1 and QE2, it has the desired long-term effect.

Taking a look at EBP, which is - roughly speaking - the component of spread between an index of rates of return on corporate securities and a similar maturity government bond rate that is left after removing the component due to default risk. Hence, as Gertler and Karadi (2015) argue, it can be interpreted as measuring the spread between yields on private versus public debt, which exists due to frictions in the financial market. It can be seen, that QE1 in 2009 can lower this premium most effectively by 1.8 pp, which stimulates corporate and private borrowing and thereby investment, consumption and GDP. However, despite of the long-term yield reduction, the EBP does not fall after QE2. An intuition for this unexpected relation is given by Gertler and Karadi (2015), who say that it could be the case that the market expectations of the path of interest rates exceed the (ANN) model-implied path. This might happen because agents fail to anticipate the decline in interest rates, which leads to a reduction in interest rates across the yield curve but little or nomovement in term premia and hence the EBP. Moreover, also the level of EBP prior to the policy intervention leads toopposing reactions, indicating that premia are effectively reduced only when being inefficiently high.

Jumping back to Figure 4, negative GDP growth values are raised through the balance sheet increase in QE1, capturing the stimulating effect at negative growth values. This increase appears promptly, as reaction to the decrease of inefficient term-premia. The same holds true for the short-run (around 12 months for QE2 and 6 months for QE3) effect of the other interventions. In the longer run, however, the loose policy during the recovery process produces lower GDP growth levels. A similar finding can be drawn from the inflation paths. Figure 5 provides further evidence, that QE1 is the only intervention, effectively increasing GDP growth by about 1 pp. Adding up all peak

Figure 5: Δ^{QE} for each Intervention



Note: This figure thows the differences between QE and no QE following the impulse responses. All variables are in percentage points.

effects, GDP is raised by 1.3 pp. It can also be seen that each measure produces a positive impulse on inflation, which is - especially having in mind the different volumes of the intervetions - most distinct for QE1. The aggregate effect amounts to 0.6 pp for all interventions.

These findings suggest that when balance sheet adjustments are conducted by the central bank, the economic state significantly influences the impact of the policy. While the program's volume is not proportional to its effect (see differences in QE2 and 3), it can even have undesireable negative effects (see QE3), depending on the depth of the crisis. One conclusion of this exercise is therefore, that the nonlinear setup (based on ANNs or other nonlinear methods) is beneficial in order to reveal these disproportionalities. Furthermore, the long-run effects on GDP and inflation are undesireable, which allows for the conclusion that a balance-sheet normalization plan is key to an efficient policy strategy. Moreover, term-premia as well as the real sector react to expected balance sheet changes promptly, which underlines the role of forward guidance by the central bank.

Comparing these findings to other literature is somehow challenging, as modeling approaches, analyzed scenarios and variable definitions vary. The GDP growth increase by 1 pp through QE1 in the counterfactual (compared to no QE), may be opposed to the crisis experiment in Gertler and Karadi (2013), which indicates a GDP increase by about 2.2 pp with QE. Inflation in their experiment is increased by 3 pp (0.9 pp here)[7] and the yield is decreased by about 20 bp (also 20 bp in this analysis)[8]. Hence, while the initial effect of QE1 on the financial sector is equivalent to their findings, the effect on the real sector captured by the ANN-based analysis reveals only half the effect found by Gertler and Karadi (2013).

4. Conclusion

This paper provides new evidence on the effects of Federal Reserve's large-scale Treasury securities purchases on the financial markets and the macroeconomy, using information from FOMC announcements as policy measure. Employing novel techniques of artificial intelligence, a system of neural networks is estimated using six representative variables - FFR, 10-year yield, EBP, inflation, GDP growth, LSAP measure - for the period between 1990 and 2015. These networks have the advantage of being universal function approximators, flexible enough to detect underlying (non)-linear interrelations in the data.

Within a historical counterfactual analysis, the effects of QE versus no intervention can be contrasted over time. While the effects on the financial sector are clearly visible, real effects are less pronounced and even become negative towards the end of the analyzed period. This can be explained through positive "on-impact" effects of balance sheet increases, and negative long-term reactions. Impulse response functions disentangle the impact of the individual policy. The results indicate heterogeneity in the transmission of the policy, as well as its impact, which depends on the economic stance of the economy prior to the intervention (the deeper the crisis, the more power). This points to disproportionalities which allow the methodological conclusion that using ANNs (or another nonlinear method) is beneficial to answer the question at hand. Economically, the results outlay that QE should rather be conducted when the economy is in a deep recession. Otherwise, effects can be marginal or even undesireable. Hence, the timing of balance sheet adjustments is particularly important. Furthermore, the volume of the intervention should be questioned, as GDP growth (inflation) is increased by around 1.3 pp (0.6 pp) only, adding up the effects of all interventions on impact. These effects are rather small, ranging at the lower bound compared to other research. Furthermore, a balance-sheet normalization strategy should be disposed in order to avoid negative long-run effects. Besides that, the results outlay that forward guidance by the central

[7] The value of 0.9 is calculated by transforming the peak effect of monthly price level changes to quarterly changes.

[8] Similarly, the monthly yield is transformed to quarterly values to get 20bp.

bank plays an important role, as it has immediate effects, and should therefore be carefully applied.

The transmission channel of QE seems to vary over time and, in this regard, future work could improve modeling by expanding the set of financial variables to include more instruments than the 10-year yield and the EBP. Furthermore, a subsequent analysis could allow for a broader balance sheet representation by the policy measure. Nevertheless, this work has shown great potential of using neural networks for unconventional policy analyses, which should be further pursued in subsequent research projects.

Acknowledgments

The paper represents the author's personal opinion and does not necessarily reflect the views of the Institute for Monetary and Financial Stability or the Goethe University. Any errors are mine. I would like to thank Natascha Hinterlang, Volker Wieland, Alexander Meyer-Gohde and further colleagues and friends for fruitful comments and discussions.

Conflicts of interest: none.

References

BAUMEISTER, C. AND L. BENATI (2010): "Unconventional monetary policy and the great recession-Estimating the impact of a compression in the yield spread at the zero lower bound," .

BLANCHARD, O., E. CERUTTI, AND L. SUMMERS (2015): "Inflation and Activity–two Explorations and their Monetary Policy Implications," Tech. rep., National Bureau of Economic Research.

CARLSTROM, C. T., T. S. FUERST, AND M. PAUSTIAN (2017): "Targeting Long Rates in a Model with Segmented Markets," *American Economic Journal: Macroeconomics*, 9, 205–42.

CHEN, H., V. CÚRDIA, AND A. FERRERO (2012): "The Macroeconomic Effects of Large-Scale Asset Purchase Programmes," *The economic journal*, 122.

CHUNG, H., J.-P. LAFORTE, D. REIFSCHNEIDER, AND J. C. WILLIAMS (2012): "Have we underestimated the likelihood and severity of zero lower bound events?" *Journal of Money, Credit and Banking*, 44, 47–82.

COIBION, O. AND Y. GORODNICHENKO (2015): "Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation," *American Economic Journal: Macroeconomics*, 7, 197–232.

DOLADO, J., R. M.-D. PEDRERO, AND F. J. RUGE-MURCIA (2004): "Nonlinear monetary policy rules: some new evidence for the US," *Studies in Nonlinear Dynamics & Econometrics*, 8.

DOLADO, J. J., R. MARIA-DOLORES, AND M. NAVEIRA (2005): "Are monetary-policy reaction functions asymmetric?: The role of nonlinearity in the Phillips curve," *European Economic Review*, 49, 485–503.

ELLISON, M. AND A. TISCHBIREK (2014): "Unconventional Government Debt Purchases as a Supplement to Conventional Monetary Policy," *Journal of Economic Dynamics and Control*, 43, 199 – 217.

FORESEE, F. D. AND M. T. HAGAN (1997): "Gauss-Newton Approximation to Bayesian Learning," in *Proceedings of International Conference on Neural Networks (ICNN'97)*, IEEE, vol. 3, 1930–1935.

GAGNON, J. (2010): *Large-Scale Asset Purchases by the Federal Reserve: Do They Work?*, Federal Reserve Bank of New York Staff Report, no. 441.

GERTLER, M. AND P. KARADI (2013): "QE 1 vs. 2 vs. 3. . . : A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool," *International Journal of Central Banking*, 9, 5–53.

——— (2015): "Monetary policy surprises, credit costs, and economic activity," *American Economic Journal: Macroeconomics*, 7, 44–76.

GILCHRIST, S. AND E. ZAKRAJŠEK (2012): "Credit Spreads and Business Cycle Fluctuations," *American economic review*, 102, 1692–1720.

GLOROT, X., A. BORDES, AND Y. BENGIO (2011): "Deep Sparse Rectifier Neural Networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, JMLR Workshop and Conference Proceedings, 315–323.

HAMILTON, J. D. AND J. C. WU (2012): "The Effectiveness of alternative Monetary Policy Tools in a Zero Lower Bound Environment," *Journal of Money, Credit and Banking*, 44, 3–46.

HANIN, B. AND M. SELLKE (2018): "Approximating Continuous Functions by ReLU Nets of Minimal Width," *arXiv:1710.11278 [cs, math, stat]*, arXiv: 1710.11278.

HORNIK, K., M. STINCHCOMBE, AND H. WHITE (1989): "Multilayer Feedforward Networks are Universal Approximators," *Neural networks*, 2, 359–366.

KIM, K., T. LAUBACH, AND M. WEI (2020): "Macroeconomic effects of large-scale asset purchases: new evidence," .

LUCAS, R. E. (1976): “Econometric policy evaluation: A critique,” in *Carnegie-Rochester conference series on public policy*, North-Holland, vol. 1, 19–46.

MACKAY, D. J. (1992): “Bayesian Interpolation,” *Neural Computation*, 4, 415–447.

NEELY, C. J. ET AL. (2010): *The Large Scale Asset Purchases had large international Effects*, Federal Reserve Bank of St. Louis, Research Division.

PRIMICERI, G. E. (2005): “Time varying structural vector autoregressions and monetary policy,” *The Review of Economic Studies*, 72, 821–852.

SMALTER HALL, A. AND T. R. COOK (2017): “Macroeconomic Indicator Forecasting with Deep Neural Networks,” *Federal Reserve Bank of Kansas City Working Paper*.

SWANSON, E. T. (2011): “Let’s twist again: A high-frequency Event-Study Analysis of Operation Twist and its Implications for QE2,” *Brookings Papers on Economic Activity*, 2011, 151–188.

TICKNOR, J. L. (2013): “A Bayesian regularized artificial neural network for stock market forecasting,” *Expert Systems with Applications*, 40, 5501–5506.

TÄNZER, A. (2021): “Multivariate Macroeconomic Forecasting: FromDSGE and BVAR to Artificial Neural Networks,” *Mimeo*.

TÄNZER, A. AND V. WIELAND (2021): “The Effectiveness of Central Bank Purchases of long-term Treasury Securities: A Comparative Approach,” *Mimeo*.

VERSTYUK, S. (2020): “Modeling Multivariate Time Series in Economics: From Auto-Regressions to Recurrent Neural Networks,” *Available at SSRN 3589337*.

Appendix

A. Data Sources and Transformation

The employed data series are mostly extracted from the FRED database. The effective federal funds rate (FEDFUNDS) - monthly, not seasonally adjusted - is transformed from year on year to month on month values, dividing all entries by twelve. A similar transformation is conducted with the 10-year Treasury constant maturity rate (DGS10) - monthly, not seasonally adjusted. Data on the consumer price index for all urban consumers (CPIAUCSL) - monthly, seasonally adjusted - are transformed to the inflation measure calculating monthly growth rates. Further, the industrial production (INDPRO) - monthly, seasonally adjusted - is transformed to monthly growth rates and represents the GDP growth measure[9]. Finally, the excess bond premium is downloaded from the Federal Reserve's webpage[10]. As stated here, the EBP is a financial indicator introduced by Gilchrist and Zakrajšek (2012), which is a component of corporate bond credit spreads. It is not directly linked to expected default risk, but provides a measure of *risk appetite* in the corporate bond market.

[9] The industrial production is an econnomic indicator that measures real output for all facilities located in the United States manufacturing, mining, electric and gas utilities.

[10] https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/updating-the-recession-risk-and-the-excess-bond-premium-20161006.html

IMFS WORKING PAPER SERIES

Recent Issues

203 / 2024	Gerhard Rösl	A present value concept for measuring welfare
202 / 2024	Reimund Mink Karl-Heinz Tödter	Staatsverschuldung und Schuldenbremse
201 / 2024	Balint Tatar Volker Wieland	Taylor Rules and the Inflation Surge: The Case of the Fed
200 / 2024	Athanasios Orphanides	Enhancing resilience with natural growth targeting
199 / 2024	Thomas Jost Reimund Mink	Central Bank Losses and Commercial Bank Profits – Unexpected and Unfair?
198 / 2024	Lion Fischer Marc Steffen Rapp Johannes Zahner	Central banks sowing the seeds for a green financial sector? NGFS membership and market reactions
197 / 2023	Tiziana Assenza Alberto Cardaci Michael Haliassos	Consumption and Account Balances in Crises: Have We Neglected Cognitive Load?
196 / 2023	Tobias Berg Rainer Haselmann Thomas Kick Sebastian Schreiber	Unintended Consequences of QE: Real Estate Prices and Financial Stability
195 / 2023	Johannes Huber Alexander Meyer-Gohde Johanna Saecker	Solving Linear DSGE Models With Structure Preserving Doubling Methods
194 / 2023	Martin Baumgärtner Johannes Zahner	Whatever it takes to understand a central banker – Embedding their words using neural networks
193 / 2023	Alexander Meyer-Gohde	Numerical Stability Analysis of Linear DSGE Models – Backward Errors, Forward Errors and Condition Numbers
192 / 2023	Otmar Issing	On the importance of Central Bank Watchers
191 / 2023	Anh H. Le	Climate Change and Carbon Policy: A Story of Optimal Green Macroprudential and Capital Flow Management
190 / 2023	Athanasios Orphanides	The Forward Guidance Trap
189 / 2023	Alexander Meyer-Gohde Mary Tzaawa-Krenzler	Sticky information and the Taylor principle

188 / 2023	Daniel Stempel Johannes Zahner	Whose Inflation Rates Matter Most? A DSGE Model and Machine Learning Approach to Monetary Policy in the Euro Area
187 / 2023	Alexander Dück Anh H. Le	Transition Risk Uncertainty and Robust Optimal Monetary Policy
186 / 2023	Gerhard Rösl Franz Seitz	Uncertainty, Politics, and Crises: The Case for Cash
185 / 2023	Andrea Gubitz Karl-Heinz Tödter Gerhard Ziebarth	Zum Problem inflationsbedingter Liquiditätsrestriktionen bei der Immobilienfinanzierung
184 / 2023	Moritz Grebe Sinem Kandemir Peter Tillmann	Uncertainty about the War in Ukraine: Measurement and Effects on the German Business Cycle
183 / 2023	Balint Tatar	Has the Reaction Function of the European Central Bank Changed Over Time?
182 / 2023	Alexander Meyer-Gohde	Solving Linear DSGE Models with Bernoulli Iterations
181 / 2023	Brian Fabo Martina Jančoková Elisabeth Kempf Luboš Pástor	Fifty Shades of QE: Robust Evidence
180 / 2023	Alexander Dück Fabio Verona	Monetary policy rules: model uncertainty meets design limits
179 / 2023	Josefine Quast Maik Wolters	The Federal Reserve’s Output Gap: The Unreliability of Real-Time Reliability Tests
178 / 2023	David Finck Peter Tillmann	The Macroeconomic Effects of Global Supply Chain Disruptions
177 / 2022	Gregor Boehl	Ensemble MCMC Sampling for Robust Bayesian Inference
176 / 2022	Michael D. Bauer Carolin Pflueger Adi Sunderam	Perceptions about Monetary Policy
175 / 2022	Alexander Meyer-Gohde Ekaterina Shabalina	Estimation and Forecasting Using Mixed-Frequency DSGE Models
174 / 2022	Alexander Meyer-Gohde Johanna Saecker	Solving linear DSGE models with Newton methods
173 /2022	Helmut Siekmann	Zur Verfassungsmäßigkeit der Veranschlagung Globaler Minderausgaben