A Nonlinear Model of Economic Data Related to the German Automobile Industry

S. Dietz[*]

Abstract Prediction of economic variables is a basic component not only for economic models, but also for many business decisions. But it is difficult to produce accurate predictions in times of economic crises, which cause nonlinear effects in the data. Such evidence appeared in the German automobile industry as a consequence of the financial crisis in 2008/09, which influenced exchange rates and automobile manufacturers' share prices. In this essay a new method of time series analysis, Autoregressive Neural Network Vector Error Correction Models (ARNN-VECM), based on the concept of nonlinear cointegration of Escribano and Mira [1] and the universal approximation property of single hidden layer feedforward neural networks of Hornik [2] is used for prediction and analysis of the relationships between 4 variables related to the German automobile industry: The US Dollar to Euro exchange rate, the industrial production of the German automobile industry, the sales of imported cars in the USA and an index of shares of German automobile manufacturing companies. The model differentiates between two kinds of relationships: The long run linear relationship (the cointegration relationship) is estimated with a 2SLS method, whereas the stock index is used as instrumental variable. This is due to the fact that share prices are an incentive for management to optimize its operating business. The short run adjustment is the nonlinear part of the model, in which the long run relationship is adjusted at nonlinear temporal occurrence. This part of the model improves the prediction power of the ARNN-VECM significantly, as it is able to handle the crisis of 2008/09. Monthly data from January 1999 to September 2009 are used for estimation of the models. They are estimated using several testing and inference methods for optimal model design as well as a customized Levenberg-Marquardt algorithm for optimization of the parameters. Prediction results are compared to various linear and nonlinear univariate and multivariate models, which are all outperformed by the ARNN VECM concerning short run prediction.

Keywords Nonlinear Time Series Analysis, Vector Error Correction, Neural Networks, Financial Crisis, German Automobile Industry.

1 Introduction

The financial crisis of 2008/09 started with the subprime crisis and gradually faced the world's financial and real economy. With the falling purchasing power of the US consumers, the German automobile industry was heavily affected by the crisis in one of its key markets with just a few months delay. Moreover with the devaluation of the US dollar compared to the Euro starting already in the mid 2000's, production costs in relation to sales prices in the US increased. The crisis has just boosted this economic exposure effect. It has to be expected, that with decreasing sales also share prices of German car manufacturers would fall and

_____

* Corresponding Author.
    E-mail: sjd@hotmail.de (S. Dietz)

S. Dietz
    University of Passau, Germany

consequently the management would be forced to act: Production costs had to be transferred from Euro to US Dollar, with the result that the industrial production of the automobile industry in Germany would decrease. In the following parts, an appropriate econometric model is used to investigate those relationships.

Economic crises have led to many doubts and discussions about econometric modeling. In particular linear methods like the autoregressive moving average processes of Box and Jenkins [3] are not always able to reflect the economic facts in data from times of crises. Such econometric models in general consist of two parts: The first one is predictable and the other one is uncertain but has to satisfy some assumptions. To overcome the inadequacy of linear methods, two approaches are possible: To weaken the conditions on the assumptions in the uncertain part, in particular on assumed distributions (nonparametric models) or to modify the predictable part such that it explains a larger part of the data (nonlinear models). The method used here belongs to this second approach. It differs from other nonlinear methods like kernel regression as it is clearly parametric and thus applicable to small data sets with 100- 200 observations like they are typical for monthly observed economic variables. As it is based on artificial neural networks, it has however the property to approximate any unknown function [2].

To cover all the reciprocally effects of a set of variables it is not only necessary to include nonlinearity in econometric models, but also to consider long and short run effects. A well-known approach to detect long run relationships is the concept of cointegration [4-5]. Two or more variables are called cointegrated if there is a long run co-movement between them. Short run effects in contrast take only effect over a finite time and balance out each other in the long run. Our model will also imply the concept of cointegration and uses therefore the nonlinear approach of Escribano and Mira [1].

This paper is structured as follows: In the second part the data and their nonlinear structure are explained. The third part introduces the econometric model we use. In the fourth part our results are compared with results using some other econometric methods and checked whether all assumptions are fulfilled. The fifth part concludes the essay.

2 The Data

The data we use here are monthly observations from January 1999 to September 2009 (total 129 observations) of the following variables:
- Industrial production of Germany: Index of the industrial production of the German car manufacturing and car parts manufacturing industry adjusted for working days with the average of 1999=100. As this series contains linear seasonal effects, it is seasonal adjusted by splitting of the seasonal part.
- Sales USA: Index of sales of imported foreign cars in the USA with the average of 1999=100. This series also contains linear seasonal effects, therefore it is seasonal adjusted by splitting of the seasonal part.
- Exchange rate USD/EUR: Dollar per 1 Euro exchange rate provided by the German federal reserve bank.
- German car manufacturer index: Prices of shares traded at Frankfurt stock exchange in Euro are used. The following shares of premium car manufacturers are included: BMW common stock, Daimler common stock and preferred stocks of the Dr. Ing. h.c. F. Porsche AG. Other car manufacturers are not considered, as their exports to the USA do not

contribute a significant part to their business. The German car manufacturer index is calculated as a Laspeyres index. The sales amount of each individual share divided by the sum of sales amount of all three companies in 1998 are used as weights. In addition, logarithms of the series are calculated to bring the series in accordance with the other ones concerning the behaviour of the first differences.
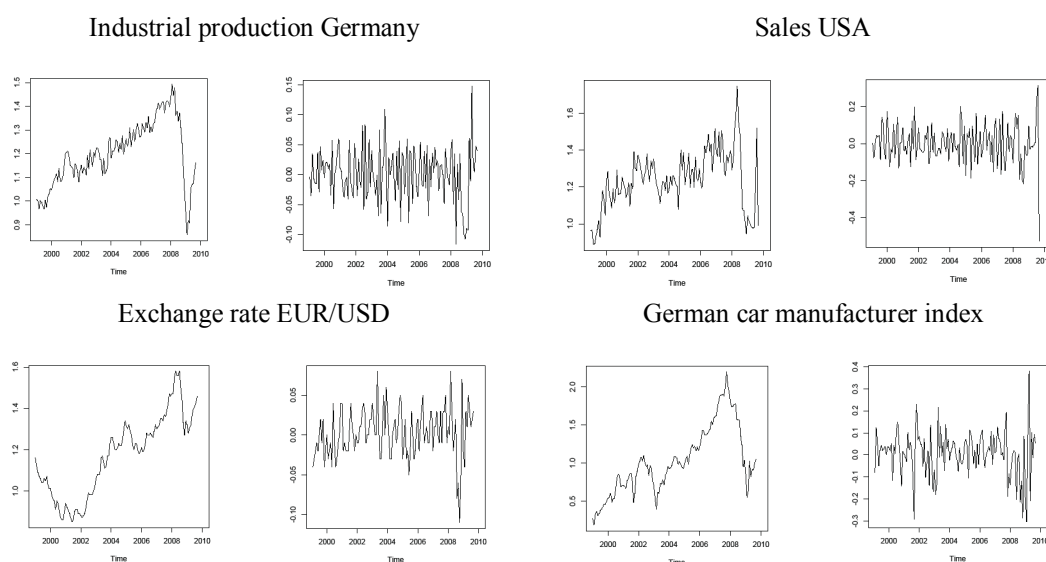
Industrial production Germany   Sales USA

Exchange rate EUR/USD   German car manufacturer index

Fig. 1 Data in levels and first differences

If the time series in fig 1 are considered, there is obvious evidence for a nonstandard behaviour in the last 10% of the observed data. This observation is due to the impact of the financial crisis. Before the model is built, it has to be examined by appropriate tests if additional data transformation is necessary. This is due to the three-step procedure (data preparation, parameter estimation and model validation/inference) proposed by Box and Jenkins [3] for time series models. First, the degree of integration of the data has to be determined for estimation of the models. To investigate this, one of the well-known unit root tests can be used, for example the augmented Dickey-Fuller test [6]. Alternatively, the rank augmented Dickey-Fuller test, proposed by Hallman [7] can be used. Such tests indicate that the data for the four variables are nonstationary, and their first differences are stationary. This means that they are integrated of order 1 (~I(1)), which is important for cointegrated models. This will further be discussed in the following section.

Furthermore, it has to be examined, if a nonlinear part would significantly improve the econometric model, or if a linear model is sufficient. The test of Teräsvirta et al. [8] can be used for this purpose: It uses a Taylor polynomial to approximate an unknown nonlinear part and examines if this reduces the residual sum of squares significantly. In our case all data series are nonlinear whereas the nonlinearity is contained mainly in the last 10% and can be intuitively explained by the influence of the crisis. Table 1 shows the results of the test of Teräsvirta et al. [8] for up to 3 lags using the data including the crisis, whereas table 2 shows the results for data without the crisis. In table 1 all test statistics are above the critical value (with the only exception exchange rate at lag 1), which is an indicator for nonlinearity. In table 2 in contrast all test statistics are below the critical value. This shows that the nonlinearity is mainly caused by the crisis of 2008/09.

Table 1 Nonlinearity test of Teräsvirta and et al. [8] for the variables including the crisis

Series	Lag	Chi-square test statistic	Critical value (95%)
Industrial	1	12.5899	5.9915
production	2	18.9509	14.0671
Germany	3	29.1732	26.2962
Sales USA	1	6.2537	5.9915
	2	22.5614	14.0671
	3	33.1896	26.2962
Exchange	1	4.7863	5.9915
rate	2	17.4252	14.0671
USD/EUR	3	39.6585	26.2962
German car	1	11.5989	5.9915
manufacturer	2	20.2478	14.0671
index	3	44.2536	26.2962

Table 2 Nonlinearity test of Teräsvirta et al. [8] for the variables without the crisis (first 100 observations, January 1999 to April 2007)

Series	Lag	Chi-square test statistic	Critical value (95%)
Industrial	1	0.4235	5.9915
production	2	8.8055	14.0671
Germany	3	24.2016	26.2962
Sales USA	1	2.4662	5.9915
	2	8.4997	14.0671
	3	14.2617	26.2962
Exchange	1	1.9647	5.9915
rate	2	7.3563	14.0671
USD/EUR	3	24.7121	26.2962
German car	1	2.3462	5.9915
manufacturer	2	11.1990	14.0671
index	3	24.3844	26.2962

3 The econometric model
3.1 Vector error correction models and nonlinearity

In this section the Autoregressive Neural Network Vector Error Correction Model should be introduced step by step. At first, we start with a linear Vector Autoregressive Model (VAR): Let $Y_t = (y_{1t},...,y_{mt})'$ be a vector of m variables representing a stochastic process depending of time t. An VAR(n) model of this process is

$$Y_t = A_1 Y_{t-1} + ... + A_n Y_{t-n} + E_t$$

where E_t is a vector of multivariate i.i.d. Gaussian distributed errors with expectation 0. This process model consists of two kinds of variables: The first n terms are called predictable part, as they can be determined by appropriate methods, whereas the last term is the stochastic or uncertain part, which accounts for unpredictable events. To avoid spurious regression within this equation Granger and Newbold [9] propose to use only stationary time series for estimation. Stationary in a general sense means that the moments of the distribution of the

time series do not change over time. If Gaussian distribution is used for the errors, stationarity of mean and covariance between the time lags (weak stationarity) is sufficient. However to stationarise the data via differentiation - like every data transformation - may bring along some loss of information. In the following we use Δ as notation for the first differences.

Engle and Granger [4] introduced the concept of cointegration to solve this problem of stationarity. The basic idea is that within a vector of nonstationary variables a stationary equilibrium exists. This equilibrium represents the long run relationship between the variables. Cointegration can technically be defined as follows: A vector of m variables integrated of order 1, Y_t, is called cointegrated if there exists a matrix C such that $C'Y_t$ is stationary. C has a rank of $(r \times m)$ with $r<m$. This condition excludes that Y_t consists of stationary variables, which would be the case if $(m=r)$. Once this long run equilibrium is identified, there are still short run effects in the data. Such short run effects may be the influence of some temporary economic phenomenon. The Vector Error Correction Model (VECM) accounts for that: Using some transformation like in Johansen [5] a linear VAR of a vector of I(1)-variables can be transformed into a VECM representation, where the long run and the short run parts are separated:

$$\Delta Y_t = \mathrm{K}C'Y_{t-1} + \mathrm{A}_1 \Delta Y_{t-1} + ... + \mathrm{A}_n \Delta Y_{t-n} + \mathrm{E}_t$$

The left hand of the equation represents the change of the variable vector from state *t-1* to state *t*. The *(r×m)* matrix K in this equation is the linear adjustment of the long run equilibrium at the present state, whereas the other terms account for the influence of short run effects from the previous states on the present state. The error term E_t is known from above. This representation is used in economics to detect the long run equations and to investigate the nature of short and long run influences.

As already discussed in section 2, the variables used here are nonlinear. Hence in the following the nonlinear VECM of Escribano and Mira [1] is introduced. It assumes that a linear long run equilibrium between the nonlinear variables exist. This has to be adjusted at the nonlinear states by a not yet specified nonlinear adjustment function mapping from the *r*-dimensional into the *m*-dimensional space, *G: r → m*. If it is inserted into the linear VECM instead of K it becomes a nonlinear VECM :

$$\Delta Y_t = G(C'Y_{t-1}) + \mathrm{A}_1 \Delta Y_{t-1} + ... + \mathrm{A}_n \Delta Y_{t-n} + \mathrm{E}_t$$

To prove that this equation has the same properties as a linear VECM Escribano and Mira [1] use the property Near Epoch Dependency (NED) on an α-mixing sequence as a more general version of the term I(0). But as far as *G* fulfills the general Lipschitz conditions, the property I(0) can be used without any additions or modifications.

3.2 The ARNN-VECM

The general nonlinear VECM of Escribano and Mira [1] can now be concretised using a general nonlinear function. First of all the cointegration relationship has to be determined. In the following we assume, that only one cointegration relationship exists (*r=1*) such that $C'Y_t$ is scalar. This is sufficient as in reality often only one cointegration relationship exists which is useful for economic reasons. We will use a special kind of artificial neural networks, which

have a universal approximation property. That means that they can approximate any function arbitrarily and precisely. Those neural networks should have a linear as well as a nonlinear part, have only one hidden layer and a bounded nonpolynomial activation function. In neural network terminology layer means a set of variables: In the 3-layer network used in the following a input layer (the scalar cointegration relationship), a output layer (the result of the nonlinear transformation, a m-dimensional vector) and a hidden layer (the nonlinear transformation via the nonlinear function $G: 1 \rightarrow m$) exist.

The hidden layer has to be investigated further: It consists of a linear part as well as of h so called neurons. Thus such a neural network can be interpreted as an augmented linear model: Because the linear part is not sufficient to explain the behaviour of the data, it is extended by h neurons. In each of them the input is at first transformed using a bias vector Γ_0 and a weight vector Γ_1, both with dimension ($m \times 1$). After this linear transformation the input has dimension ($m \times 1$), the dimension of the output. Now the nonlinear transformation is performed using the activation function $\Psi: m \rightarrow m$, which is bounded and nonpolynomial. In the artificial neural network literature (e.g. [10]) often so called sigmoid (because of their S-like plot) functions like the tangens hyperbolicus or the logistic function are used, but of course others like the cosine or a threshold function are possible. Finally the results of the neurons are weighted by a scalar β and summarized with the linear part. The activation function is the same for all h neurons, whereas the weights Γ_0, Γ_1, and β have individual values which can be estimated using a nonlinear optimization algorithm. The neural network used here should be called Autoregressive Neural Network (ARNN) as the time lag Y_{t-1} is included (the cointegration relationship). To complete the nonlinear VECM, the function $G(C'Y_t)$ can be rewritten as ARNN:

$$G(C'Y_{t-1}) = \begin{pmatrix} \alpha_{01} \\ \vdots \\ \alpha_{0m} \end{pmatrix} + \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{1m} \end{pmatrix} C'Y_{t-1} + \sum_{j=1}^{h} \Psi\left(\begin{pmatrix} \gamma_{01j} \\ \vdots \\ \gamma_{0mj} \end{pmatrix} + \begin{pmatrix} \gamma_{11j} \\ \vdots \\ \gamma_{1mj} \end{pmatrix} C'Y_{t-1} \right) \beta_j$$

Hornik [11], based on Hornik, Stinchcombe and White [2], shows that an ARNN as used here can approximate any function if there are sufficient neurons. This property can intuitively be explained by the fact that the output of each neuron in the nonlinear part is a constant which depends on a nonlinear adjustment of the cointegration relationship. If a sufficient large number of constants is included in the equation, the constants adjust at nonlinearities like structural breaks or multiplicative effects in the original data. Hence, ARNN are a simple structured method to approximate nonlinear autoregressive processes, which gets more neurons by getting more preciseness and thus the parameters are included.

4 Results
4.1 Estimation of the ARNN-VECM

To validate the estimation results in the following, the original data set is partitioned into an estimation and a validation subset. The validation subset is used in the following to evaluate prediction results from the estimated model. Here the estimation subset consists of 120 values, the validation subset of 8 values (from the original T=129 1 observation gets lost by calculating the first differences).

The ARNN-VECM is estimated in two steps: At first the linear long run equilibrium, the cointegration relationship, is calculated. In the second step the ARNN-VECM which surrounds the cointegration relationship is estimated. For calculation of the cointegration relationship the 2SLS method is used. OLS estimation would be inappropriate here, because spurious regression may appear if nonstationary variables are directly regressed on each other. The regression here is performed in two steps: First of all the sales USA are regressed on exchange rate and the German car manufacturer index. In the second step the industrial production in Germany is regressed on the estimated sales USA from the first step and the exchange rate. The following stationary long run relationship results:

Industrial production = 1.1515 Sales USA -0.1867 Exchange rate

The German car manufacturer index is used as structural variable here, which means that it only indirectly flows into the equation. The economic justification for this is the fact, that share prices are a management incentive, which reacts on decreasing sales and unfavorable developments of exchange rates (increasing amount of USD per 1 EUR) by falling prices. If such developments take place in the long run, share prices force the management of German car manufacturing companies to shift production into the USD currency area. Such evidence did indeed happen in 2009/10 (BMW announced to invest 750 mio EUR in its US plant and Daimler planned to shift parts of the production of the Mercedes C-class to the USA). The two parameters in the long run equilibrium can be interpreted as follows: The correlation between industrial production in Germany and sales USA is positive, which means that an increase in sales leads to an increase in production. In contrast the correlation of industrial production in Germany and the exchange rate is negative. This shows that an increasing USD (falling exchange rate) has a positive influence on industrial production. In addition the coefficient of the exchange rate is much lower than that of sales USA. This is realistic as the influence of sales USA on the industrial production in Germany is more direct and therefore stronger. The augmented Dickey-Fuller test indicates that the cointegration relationship is weakly stationary for several lags.

For further estimation of the ARNN-VECM for the 3 variables the number of lags n has to be determined. To this end the procedure of Rech and et al. [12] can be used. Therefore the variables are considered separately. For each variable for increasing lag orders a nonlinear AR process is estimated using a Taylor polynomial. For those nonlinear models the AIC is calculated. The model with the lowest AIC indicates the optimal lag order. The average optimal lag order from the separate models should be the lag order of the multivariate model. For the variables used here an optimal lag order of $n=3$ resulted from that procedure.

The second step of identification of the ARNN-VECM is to estimate the parameters of the full representation. For this purpose numerical optimisation algorithms like Newton's method, Quasi-Newton methods or genetic algorithms can be used. For the results in the following the Levenberg-Marquardt algorithm is used. This is an iterative Gauss-Newton method, which is as powerful as Newton's method but much less complex. For estimation the multivariate ARNN-VECM equation was split into 3 univariate equations. This can be done without getting in trouble with the dimensions at least because of the singular β. However the β-weights have to be pre-determined as they can be included in the estimation algorithm. The estimation procedure itself was performed starting with arbitrary determined weights (here uniformly set to 0.005) iteratively into the direction of a minimal residual sum of squares of

the ARNN-VECM. For details about that estimation procedure see Dietz [13]. In table 3 the results for a linear VAR (no hidden neuron, estimated by the OLS method) as well as ARNN-VECM with neurons $h=0$ (corresponding to a linear VECM) and $h=1,2,12$ with in-sample and 1- and 8-step out-of sample predictions compared to the validation subset (first differences) are shown. As a measure for the out-of-sample prediction Theil's inequality coefficient (THEIL) is used, which compares the predicted results to a naive forecast, which is here the arithmetic mean of the time series (only of the estimation subset). A THEIL >1 means that the naive forecast is a better predictor than the estimated result. For a series x_t of length T the THEIL is calculated for a k-step prediction by:

$$THEIL = \sqrt{\frac{\sum_{t=T+1}^{k}(x_t - \hat{x}_t)^2}{\sum_{t=T+1}^{k}(x_t - \frac{1}{T}\sum_{i=1}^{T}x_i)^2}}$$

Table 3 Estimated results for a ARNN-VECM with varying h

Series	In-sample RSME	1-step prediction THEIL	8-step prediction THEIL
VAR (27 parameters)			
Industrial production Germany	0.0396	1.2745	1.2132
Sales USA	0.0857	2.1523	1.0215
Exchange rate	0.0279	0.0387	1.1022
ARNN-VECM with $h=0$: linear VECM (36 parameters)			
Industrial production Germany	0.0369	0.7721	0.9657
Sales USA	0.0813	0.1902	1.0081
Exchange rate	0.0277	0.7210	1.1562
ARNN-VECM with $h=1$ (43 parameters)			
Industrial production Germany	0.0411	0.7284	0.7485
Sales USA	0.0811	0.2114	1.0267
Exchange rate	0.0277	4.6887	6.1888
ARNN-VECM with $h=2$ (56 parameters)			
Industrial production Germany	0.0367	1.3040	0.9049
Sales USA	0.0813	0.4615	1.0128
Exchange rate	0.0272	0.7065	0.8326
ARNN-VECM with $h=12$ (120 parameters)			
Industrial production Germany	0.0374	0.7941	0.9422
Sales USA	0.0780	0.1766	1.0138
Exchange rate	0.0280	0.7403	1.1453

To outline the increasing complexity of the ARNN-VECMs their parameters are noted in brackets in table 3. The in-sample results from table 3 show that there is not much difference between the models. However, the model with most neurons does not deliver the best prediction results. This is due to the fact that ARNN-VECMs tend to overfitting if too many parameters are included. Concerning the prediction results two things can be found out: At first including the cointegration relationship improves the out-of-sample results. Second, this improvement in the long run holds only for the industrial production time series. The fact that the out-of-sample 1-step prediction quality concerning the exchange rate time series is not improved or even gets worse lets us assume, that the exchange rate is a weak exogenous variable, which means that it exercises some influence on the cointegration relationship but is not affected by it. We also see that the number of hidden neurons has a different influence on the prediction results. Thus, the number of neurons depends on the variable which should be predicted. This is the price to pay for the flexibility of ARNN-VECM, and it is in particular due to the dimensionality.

4.2 Comparison to other methods and model validation

Finally, we want to compare the ARNN-VECM prediction results with the prediction results from some ARMA and univariate nonlinear methods. Table 4 shows such evidence with a linear ARMA model, a local linear model (a kernel regression model) and a univariate ARNN as proposed in Dietz [13]. Concerning the 8-step prediction the industrial production series as well as the exchange rate series can be better predicted using the ARNN-VECM because of the additional information from the other time series. Sales USA can be slightly better predicted by a univariate ARNN, but in the short run also the ARNN-VECM performs best.

Table 4 Results for alternative models

Series	Lags	In-sample RSME	1-step prediction THEIL	8-step prediction THEIL
ARMA				
Industrial	(2,5)	0.0399	0.5559	1.4087
Sales USA	(2,5)	0.0860	3.5896	1.0202
Exchange rate	(3,5)	0.0263	1.9766	0.8596
LLAR				
Industrial production	2	0.0445	1.4831	1.0829
Sales USA	2	0.0942	0.3769	0.9990
Exchange rate	3	0.0303	2.0056	0.9101
Univariate ARNN with h=4				
Industrial production	2	0.0378	0.4084	0.9251
Sales USA	2	0.0884	0.5777	0.9994
Exchange rate	3	0.0272	0.5032	0.8764

Finally, it should shortly be examined if the assumptions on the errors (i.i.d. with Gaussian white noise with zero mean) are fulfilled. This is an important criterion for the

validity of the model as linear models do often not fulfill those assumptions and gave rise to various extensions on the assumptions, which are much more difficult to handle (e.g. T-distribution). The independence assumption, which can be tested e.g. by the Box-test, is mostly fulfilled, but not the assumption on the distribution. Hence, we investigate this topic here: The Jarque-Bera test on nonlinearity is performed for the individual series' ARNN-VECM residuals. This test is based on the third and fourth moment of the distributions (skewness and kurtosis) and investigates weather a Gaussian distribution is present. Table 5 shows the results: In the first model the exchange rate series' residuals are not Gaussian distributed (test statistic is above the critical value) thus this model is not appropriate. For the other two models the residuals are Gaussian distributed.

Table 5 Jarque-Bera test results for the residuals of a ARNN-VECM with h=2

Series	Chi-square test statistic	Critical value (95%)
ARNN-VECM with h=1		
Industrial production	1.5965	
Sales USA	0.1232	
Exchange rate	**745.3444**	
ARNN-VECM with h=2		
Industrial production	3.7363	5,99
Sales USA	1.7951	
Exchange rate	5.7552	
ARNN-VECM with h=12		
Industrial production	1.7114	
Sales USA	1.7951	
Exchange rate	5.7552	

5 Conclusion

ARNN-VECMS are an alternative methods for analysis and prediction of data under the influence of economic crises. They provide a flexible toolkit to integrate long- and short run effects into a multivariate econometric model and to combine that with nonlinearity. If some linearly cointegrated series contain nonlinear influences, such nonlinear error correction is essential. The nonlinear error correction model as proposed here used together with a structural equation system for the cointegration relationship is especially useful for analysis of supply-demand relationships. The data used here can be seen as such a system, whereas the industrial production of car and car parts manufacturers in Germany is the supply side and the sales of imported foreign cars in the USA represents the demand side. However, the flexibility

of the models needs careful handling: Attention has to be paid to the number of parameters and the assumptions on the model.

References

1. Escribano, A., Mira, S., (2002). Nonlinear Error Correction Models. Journal of Time Series Analysis, 23(5), 509-522.
2. Hornik, K., Stinchcombe, M., White, H., (1989). Multilayer Feedforward Networks are Universal Approximators. Neural Networks, 2(5), 359-366.
3. Box, G. E. P., Jenkins, G. M., (1976). Time Series Analysis - Forecasting and Control. 2nd edition, Holden Day, San Francisco.
4. Engle, R. F., Granger, C. W., (1987). Co-Integration and Error Correction: Representation, Estimation and Testing. Econometrica, 55(2), 251-276.
5. Johansen, S., (1995). Likelihood-Based Inferecne in Cointegrated Vector Autoregressive Models. Oxford Unoiversity Press, Oxford.
6. Dickey, D. A., Fuller, W. A., (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association, 74(366), 427-431.
7. Hallman, J. J., (1990). Nonlinear Integrated Series, Cointegration and Application. PhD dissertation University of California San Diego
8. Teräsvirta, T., Lin, C. F., Granger, C. W., (1993). Power of the Neural Network Linearity Test. Journal of Time Series Analysis, 14(2), 209-220.
9. Granger, C. W., Newbold, P., (1974). Spurious Regression in Econometrics. Journal of Econometrics, 2(2), 110-120.
10. Bishop, C. M., (1995). Neural Networks for Pattern Recognition. Claredon Press, Oxford.
11. Hornik, K., (1993). Some New Results on Neural Network Approximation. Neural Networks, 6(8), 1069-1072.
12. Rech, G., Teräsvirta, T., Tschernig, R., (2001). A Simple Variable Selection Technique for Nonlinear Models. Communications in Statistics, Theory and Methods, 30(6), 1227-1241.
13. Dietz, S., (2011). Autoregressive Neural Network Processes - Univariate, Multivariate and Cointegrated Models with Application to the German Automobile Industry. PhD Dissertation University of Passau.