UNIVERSITY OF ECONOMICS IN PRAGUE

FACULTY OF INFORMATICS AND STATISTICS

BACHELOR THESIS

2013 Dmitry Borodin

UNIVERSITY OF ECONOMICS IN PRAGUE

FACULTY OF INFORMATICS AND STATISTICS

VAR Analysis of Exchange Rate Pass-Through Effect in Czech Republic

Author: **Dmitry Borodin**

Department: **Department of Econometrics**

Field of study: **Mathematical Methods in Economics**

Supervisor: **Mgr. Ing. Viktor Chrobok**

Declaration:

I hereby declare that this bachelor thesis is completely my own work and that I used only the cited sources.

Prohlašuji, že jsem bakalářskou práci na téma „VAR Analysis of Exchange Rate Pass-Through Effect in Czech Republic“ zpracoval samostatně. Veškerou použitou literaturu a další podkladové materiály uvádím v seznamu použité literatury.

V Praze dne 1. ledna 2013 ....

Dmitry Borodin

Acknowledgments:

I would like to express my feelings of gratitude to Mgr. Ing. Viktor Chrobok for his patience and incredibly useful comments, which he was always willing to share.

Abstrakt

Název práce: VAR Analýza Exchange Rate Pass-Through v České Republice
Autor: Dmitry Borodin
Katedra: Katedra ekonometrie
Vedoucí práce: Mgr. Ing. Viktor Chrobok

Cílem dané práce je empirická analýza dopadu změn kurzu České koruny na domácí úroveň cen. Tomuto se v cizí literatuře říká „exchange rate pass-through". Veškerá analýza je provedena pomoci vektorové autoregrese, na jejímž základě je odvozena funkce odezvy. Funkce odezvy umožňuje zjistit jak sílu, tak i rychlost dopadu změn kurzu koruny na úroveň cen. Celá teoretická část je věnovaná vybraným kapitolám z teorií vektorové autoregrese. Jejím úkolem bude snaha poskytnout jasný, ale na druhou stranu i komplexní pohled na VAR modely. Praktická část se potom zabývá modelováním „exchange rate pass-through".

Klíčová slova: VAR modely, funkce odezvy, exchange rate pass-through, inflace.

Abstract

Title: VAR Analysis of Exchange Rate Pass-Through Effect in Czech Republic
Author: Dmitry Borodin
Department: Department of Econometrics
Supervisor: Mgr. Ing. Viktor Chrobok

The paper will empirically investigate the strength and the speed of the exchange rate pass-through effect in the Czech Republic, i.e. the change in the domestic prices, originally caused by the volatility of the exchange rate. VAR modelling framework has been chosen as a main instrument of analysis. Vector autoregression will also be the subject of the theoretical part of the paper, which aims to provide a clear and at the same time many-sided discussion on the relevant topics. Practical part will be completely devoted to the modelling of the exchange rate pass-though.

Keywords: VAR models, impulse response analysis, exchange rate pass-through, inflation.

Table of Contents

1 Introduction

This paper is divided into two main chapters: a review of a vector autoregression theory and application of some of the VAR methods.

First chapter aims to provide a better understanding of the vector autoregressive models. We start with difference equations, as they are considered to be the mathematical basis for the later VARs. Mainly all of the formulas are at least explained or even proved. Such approach is believed to help later to avoid mechanical use of the formulas. Then we continue with the most important properties of VARs. For example, different representations of the vector autoregressive processes, stability conditions, Granger causality and impulse response functions. Some of the topics of the discussion are rarely met in the literature on VARs, despite their tremendous significance. For example, transformation of a higher order VAR(p) into the first order VAR(1) is extremely useful, as it automatically expands all the VAR(1) theory for the higher order processes. However, this topic is considered in just a few works. Generally first chapter is based on several books written both in the Czech and English languages.

Second chapter is devoted to the analysis of the exchange rate pass-through in the Czech Republic with the help of some of the vector autoregressive methods earlier mentioned in the theoretical part. Exchange rate pass-through can be defined as the change in the domestic inflation originally caused by the volatility of the exchange rates. We, firstly, provide some notes on the exchange rate pass-through and factors that might determine its extent and speed. Then we continue with a brief summary of the papers on a given subject. Generally, this topic is vastly covered by the researchers of the central banks of different countries. This is no surprise, as exchange rate pass-through (later just ERPT) is one of the main concerns of national banks, as it affects countries inflation and monetary policy.

After that we come to the descriptions of the data used in the model, model's specification, transformations and tests. Please note that the detailed output is provided only in the corresponding appendix. It is done so in order not to confuse reader with the discursive jumping around the topics. Actual analysis of the pass-through effect is carried out using accumulated impulse response functions. Following steps are later repeated on the alternative model specification with the possibility of studying so called distribution chain. We find that the results to some extent differ across the specifications of the models. However, both of the models find a modest ERPT to the consumer prices. It is in line with the conclusions of other papers on the given subject. For all of the calculation purposes EViews software is used.

2 Vector Autoregression

Due to the fact that Vector Autoregression process (VAR) can be defined as a stochastic difference equation and as technics applied on the difference equations often find their application concerning VARs, the first theoretical part will be devoted to the difference equations, which will help to get a better understanding of the later methods and of the background of the VAR modeling.

Moreover, time series econometrics heavily relies on the estimation of difference equations, containing stochastic components. Below please find some basic mathematical concepts, which are essential for time series modeling.

2.1 Difference Equations

Let's say we have a function of $y = f(t)$. If we evaluate the function at the time, when t takes specific value t^*, we will get a specific value for the dependent variable $y_{t*} = f(t^*)$. When t takes specific value t^*+h, we will similarly get $y_{t*+h} = f(t^* + h)$. First difference of y is

$$\Delta y_{t*+h} = f(t^* + h) - f(t^*) = y_{t*+h} - y_{t*}$$

The first difference of y is the difference between the values of the function evaluated at times t^*+h and t^*.

It is useful to "normalize" h so that it represents a unit change in t. For example, in financial econometrics t is used as a measure of time and h represents the length of a period. We can form a sequence of first differences:

$$\begin{aligned} \Delta y_t &= f(t) - f(t-1) = y_t - y_{t-1} \\ \Delta y_{t+1} &= f(t+1) - f(t) = y_{t+1} - y_t \\ &\dots \end{aligned}$$

Analogously sequence of the second differences is represented by:

$$\begin{aligned} \Delta^2 y_t &= \Delta(\Delta y_t) = \Delta(y_t - y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2} \\ \Delta^2 y_{t+1} &= \Delta(\Delta y_{t+1}) = \Delta(y_{t+1} - y_t) = (y_{t+1} - y_t) - (y_t - y_{t-1}) = y_{t+1} - 2y_t + y_{t-1} \\ &\dots \end{aligned}$$

No higher order differences are usually used in time series analysis. Suppose, we have an autoregression process

$$y_t = a_0 + \sum_{i=1}^{n} a_i y_{t-i} + \varepsilon_t$$

It can be easily represented in the terms of the difference operator, if we subtract y_{t-1}:

$$\Delta y_t = y_t - y_{t-1} = a_0 + (a_1 - 1)y_{t-1} + \sum_{i=2}^{n} a_i y_{t-i} + \varepsilon_t$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=2}^{n} a_i y_{t-i} + \varepsilon_t$$

There is no doubt that the later equation is just a modification of the original autoregression process.

A solution of a difference equation is a function, where y_t is explained with the ε_t sequence, t and some initial condition of y_t sequence. An important property of the solution is that it satisfies the difference equation for all the values of t and ε_t. If we substitute the solution into the difference equation, we will end up with an identity. Solution is also rarely unique.

There are several ways of finding a solution for a difference equation. Firstly, it is possible to apply an iteration method, which is simply based on a successive forward (or backward) iteration of the y sequence. Regardless the fact that iteration method is time-consuming, it is considered to be relatively easy for the understanding.

Say we have a process y_t, described by a following equation:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

y_1 must follow:

$$y_1 = a_0 + a_1 y_0 + \varepsilon_1$$

y_2 will be given by:

$$\begin{aligned} y_2 &= a_0 + a_1 y_1 + \varepsilon_2 \\ &= a_0 + a_1(a_0 + a_1 y_0 + \varepsilon_1) + \varepsilon_2 \\ &= a_0 + a_1 a_0 + {a_1}^2 y_0 + a_1 \varepsilon_1 + \varepsilon_2 \end{aligned}$$

And y_3 is:

$$y_3 = a_0 + a_1 a_0 + {a_1}^2 a_0 + {a_1}^3 y_0 + {a_1}^2 \varepsilon_1 + a_1 \varepsilon_2 + \varepsilon_3$$

We can similarly obtain the results for y_4, ... Now it is visually clear, that for $t > 0$, we get a following solution:

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + {a_1}^t y_0 + \sum_{i=0}^{t-1} a_1^i \varepsilon_{t-i}$$

Quite an important complication is the fact that we may not be provided with the value of the initial condition y_0. We can no longer iterate neither forward, nor backward till $t = t_0$, since the previous equation will not be a solution, as the value of y_0 is not known.

But if we iterate "below" $t = t_0$, we can substitute y_0 with $a_0 + a_1 y_{-1} + \varepsilon_0$. Continuing to iterate back another *m* periods, we obtain:

$$y_t = a_0 \sum_{i=0}^{t+m} a_1^i + {a_1}^{t+m+1} y_{-1-m} + \sum_{i=0}^{t+m} a_1^i \varepsilon_{t-i}$$

If we assume that $|a_1| < 1$, then the infinite sum $\sum_{i=0}^{t+m} a_1^i \to 1/(1 - a_1)$ and ${a_1}^{t+m+1}$ converges to zero. We get:

$$y_t = a_0/(1 - a_1) \, + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}$$

We can assure ourselves that this is a solution by substituting it into the original differential equation. What we will receive is an identity.

$$\frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} = a_0 + a_1 \left[a_0/(1 - a_1) \, + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-1-i} \right] + \varepsilon_t$$

$$\frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} = \frac{a_0(1 - a_1) - a_1 a_0}{(1 - a_1)} + a_1 \left[\sum_{i=0}^{\infty} a_1^i \varepsilon_{t-1-i} \right] + \varepsilon_t$$

Please note that it is not a unique solution.

If $|a_1| > 1$ a possible way of getting to the solution is to iterate forward, as ${a_1}^{t+m}$ does not converge to zero, the previous method is no longer appropriate:

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + {a_1}^t y_0 + \sum_{i=0}^{t-1} a_1 \varepsilon_{t-i}$$

In this case we still need to know the value of the initial condition y_0. Values of y_t may increase exponentially in the absolute values, solution will not be stable. Please note, that when $|a_1| \geq 1$ the effect of all disturbances does not disappear over time, in other words each disturbance has a permanent effect on y_t.

For higher order processes it is better to use another methodology. Earlier we have been looking for a particular solution of a differential equation. What we need now is to find a general solution. It is a homogenous solution(s) plus particular solution.

In the case of the n-th order equation, we have to find *n* homogenous solutions. We can obtain particular solution similarly to what we have done before. However, it might be a bit more numerically difficult.

Consider the following example, in which we will be looking for homogenous solutions:

$$y_t - a_1 y_{t-1} - a_2 y_{t-2} = 0$$

Homogenous solution must have a form of $y_t^h = A\alpha^t$ (h stands for homogenous, A is an arbitrary constant). It is a solution to a homogenous equation, which is an equation, where forcing function is zero. If we substitute it into the previous equation, we get:

$$A\alpha^t - a_1 A\alpha^{t-1} - a_2 A\alpha^{t-2} = 0$$

Then we should divide the later equation by $A\alpha^{t-2}$:

$$\alpha^2 - a_1\alpha^1 - a_2 = 0$$

This equation is called characteristic. Solving characteristic equation we get two values of α: characteristic roots. Each of them represents a homogenous solution for the initial problem. Those solutions however are not unique. In fact, any linear combination of them is also a solution. Complete homogenous solution has a form of:

$$y_t^h = A_1\alpha_1^t + A_2\alpha_2^t$$

Since $y_t^h = A\alpha^t$ holds, it is clear that if $|\alpha| > 1$, y_t^h will explode.

One of the most convenient ways of describing the stability condition is to note that the characteristic roots whether real or complex must lie within the unit circle.

Such a method can be successfully used regarding higher order difference equations. In the case of the *n*th order polynomial, we get *n* solutions for α. Characteristic roots again might be real or complex. Stability condition is also the same.

In order to find particular solutions it is often very convenient to use lag operators. Lag operator is defined as:

$$L^i y_t = y_{t-i}$$

Lag operator addresses to y_t lagged by *i* periods. It is also worth mentioning, that lag operators may provide more elegant way of describing *p*-th order equation:

$$y_t\left(1 - a_1 L - a_2 L^2 - \cdots - a_p L^p\right) = y_t A(L) = a_0 + \varepsilon_t$$

The use of lag operators while looking for a particular solution is following:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t = a_0 + a_1 L y_t + \varepsilon_t$$

$$= (a_0 + \varepsilon_t)/(1 - a_1 L)$$

Because $Lc = c$, where c is a constant, $\frac{a_0}{1-a_1 L} = \frac{a_0}{1-a_1} = a_0 + a_1 a_0 + a_1^2 a_0 + \cdots$ and lagging $\frac{\varepsilon_t}{1-a_1 L}$, we get :

$$y_t = a_0/(1 - a_1) + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}$$

With a slight difference this method can be used both for $|\alpha| \geq 1$ and $|\alpha| < 1$. Alternatively, we can use a method of undetermined coefficients.

2.2 Introduction to Vector Autoregression

Broadly speaking, vector autoregression is a system of equations where dependent variables are regressed on lagged observations of all the variables. In other words, the future value of each of the processes is a weighted sum of past (or maybe present) values plus noise. There is also a possibility of the expansion of the model to include deterministic time trend and other exogenous variables.

In the case of VAR processes historical data determines the contribution of each of the variables, instead of automatically relying on the economic theory. However, researchers still need to address to the theory, as it might suggest what variables to include into the model and how many lags should they have.

It can be shown, that every stationary, nondeterministic process can be approximated by a VAR process. Sim's idea that all the variables should be treated without splitting them into exogenous and endogenous, distinguishes VAR from the multiple simultaneous equations. Quite significant VAR characteristic, which simplifies the use of the model, is that the system can be estimated independently equation by equation using ordinary least squares

(OLS) technique, as long as there are identical regressors in each of the equations. It is also assumed that the white noise terms are independent of the history of $\boldsymbol{y_t}$.

Matrix form of VAR(1) is following:

$$\boldsymbol{y_t} = \boldsymbol{\varphi_0} + \boldsymbol{\Phi y_{t-1}} + \boldsymbol{\varepsilon_t}$$

Where $\boldsymbol{\varphi_0}$ is a k-dimensioal constant vector, $\boldsymbol{\Phi}$ is a $k \times k$ coefficient matrix and $\boldsymbol{\varepsilon_t}$ is a vector of stochastic error terms (also called vector of innovations) with zero mean and covariance matrix $\boldsymbol{\Sigma}$.

VAR models have two dimensions: the highest lag p and the number of the endogenous variables k (or simply equations) in the model. If the VAR(1) model is extended to have more lags, then the resulting k-dimensional VAR(p) can be written in a following manner:

$$\boldsymbol{y_t} = \boldsymbol{\varphi_0} + \boldsymbol{\Phi_1 y_{t-1}} + \boldsymbol{\Phi_2 y_{t-2}} + \cdots + \boldsymbol{\Phi_p y_{t-p}} + \boldsymbol{\varepsilon_t}$$

Where $\boldsymbol{\varphi_0}$ is a $k \times 1$ vector of constants,

$\boldsymbol{\Phi_i}$ is a $k \times k$ matrix of the coefficients of endogenous variables lagged by i periods,

$\boldsymbol{\varepsilon_t}$ is a $k \times 1$ vector of the normally distributed shocks,

$\boldsymbol{y_t}$ is a $k \times 1$ vector of current or lagged values of the endogenous variables of the model,

k is a number of endogenous variables (equations) in the model.

A given model is in the unrestricted reduced form. Alternatively, VAR(1) may be represented as:

$$y_{1t} = \varphi_1 + \Phi_{11} y_{1,t-1} + \Phi_{12} y_{2,t-1} + \varepsilon_{1t},$$

$$y_{2t} = \varphi_2 + \Phi_{21} y_{1,t-1} + \Phi_{22} y_{2,t-1} + \varepsilon_{2t},$$

Where both y_{1t} and y_{2t} are supposed to be stationary, ε_{1t} and ε_{2t} are uncorrelated white noise disturbances. The following model is a first order vector autoregression, since the longest length of its lag is one. Φ_{12} represents a linear dependence of y_{1t} on $y_{2,t-1}$. Thus, if $\Phi_{12} = 0$, then y_{1t} depends only on its past values. Analogously, if $\Phi_{21} = 0$, then y_{2t} does not depend on $y_{1,t-1}$ when $y_{2,t-1}$ is known. If both of Φ_{12} and Φ_{21} equal to zero, then y_{1t} and y_{2t} series are not connected. The concurrent linear relationship between y_{1t} and y_{2t} can be measured by the off-diagonal elements of the covariance matrix Σ.

It is now useful to define the positive-definite variance matrix of ε_{1t} and ε_{2t}:

$$\sum = \begin{bmatrix} var(\varepsilon_{1t}) & cov(\varepsilon_{1t}, \varepsilon_{2t}) \\ cov(\varepsilon_{1t}, \varepsilon_{2t}) & var(\varepsilon_{2t}) \end{bmatrix}$$

A few assumptions are to be made about $\boldsymbol{\varepsilon_t}$: $\boldsymbol{\varepsilon_t} \sim N(0, \Sigma)$ is considered to be Gaussian white noise with a zero mean and a constant variance, ε_s and ε_t are independent for $t \neq s$. Under this assumptions, $y_t, \ldots, y_{t+h}$ have multivariate normal distribution.

Form of VAR(1) model mentioned earlier in the paper is often called a reduced-form model, as it does not explicitly show the concurrent dependence between the component series. It is also possible to explicitly determine the concurrent dependence between the y_{1t} and y_{2t}. For these purposes it is useful to address to the Cholesky decomposition.

2.3 Cholesky Decomposition

Firstly, lower triangular matrix should be defined as (theoretically, it can be both lower and upper):

$$\boldsymbol{L} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ l_{2,1} & 1 & & 0 & 0 \\ \vdots & & \ddots & \vdots & \\ l_{k,1} & l_{k,2} & \cdots & l_{k,k-1} & 1 \end{bmatrix}$$

For any symmetric matrix, say $\boldsymbol{A}$, exists a lower triangular matrix $\boldsymbol{L}$ (it's diagonal elements are equal to 1) and a diagonal matrix $\boldsymbol{G}$ so that $\boldsymbol{A} = \boldsymbol{LGL'}$ holds. At this point, it is worth noticing, that if $\boldsymbol{A}$ is a positive definite matrix, then the diagonal elements are also positive. Moreover, as $(\boldsymbol{AB})' = \boldsymbol{B'A'}$, then:

$$\boldsymbol{A} = \boldsymbol{L}\sqrt{\boldsymbol{G}}\sqrt{\boldsymbol{G}}\boldsymbol{L'} = (\boldsymbol{L}\sqrt{\boldsymbol{G}})(\boldsymbol{L}\sqrt{\boldsymbol{G}})'$$

This is called Cholesky decomposition; it shows that a positive-definite matrix can be diagonalized. Using the later notation:

$$\boldsymbol{G} = \boldsymbol{L^{-1}A}(\boldsymbol{L^{-1}})'$$

2.4 Structural Form of VAR

Returning to the previous discussion on the concurrent dependence between the y_{1t} and y_{2t}, as $\boldsymbol{\Sigma}$ is a positive-definite, then $\boldsymbol{\Sigma} = \boldsymbol{LGL'}$ holds, where $\boldsymbol{L}$ is a lower triangular matrix with unit diagonal elements and $\boldsymbol{G}$ is a diagonal matrix. Also $\boldsymbol{G} = \boldsymbol{L^{-1}\Sigma}(\boldsymbol{L^{-1}})'$ holds.

Introducing $\boldsymbol{b_t} = \boldsymbol{L^{-1}\varepsilon_t}$. As $E(\boldsymbol{\varepsilon_t}) = 0$, then $E(\boldsymbol{b_t}) = \boldsymbol{L^{-1}\varepsilon_t} = \boldsymbol{0}$ and due to the Cholesky decomposition $Cov(\boldsymbol{b_t}) = \boldsymbol{L^{-1}\Sigma}(\boldsymbol{L^{-1}})' = \boldsymbol{G}$.

As already have been mentioned $\boldsymbol{G}$ is a diagonal matrix. It implies that the components of $\boldsymbol{b_t}$ are not correlated.

If we multiply VAR(1) model by $\boldsymbol{L^{-1}}$ from the left, we get:

$$\boldsymbol{L^{-1} y_t} = \boldsymbol{L^{-1}}\ \boldsymbol{\varphi_0} + \boldsymbol{L^{-1} \Phi y_{t-1}} + \boldsymbol{L^{-1} \varepsilon_t}$$

$$= \boldsymbol{\varphi_0}^* + \boldsymbol{\Phi}^* \boldsymbol{y_{t-1}} + \boldsymbol{b_t}$$

Where $\boldsymbol{\varphi_0}^* = \boldsymbol{L^{-1}}\ \boldsymbol{\varphi_0}$ and $\boldsymbol{\Phi}^* = \boldsymbol{L^{-1}\Phi}$.

The k-th equation is going to look as following (if the denote the last row of $\boldsymbol{L^{-1}}$ matrix as $(l_{k,1}, \dots, l_{k,k-1}, 1)$):

$$y_{kt} + \sum_{i=1}^{k-1} l_{ki}\, y_{it} = \varphi_{k0}{}^* + \sum_{i=1}^{k} \Phi_{ki}{}^* y_{i,t-1} + b_{kt}$$

This formula represents the concurrent linear dependence of y_{kt} on y_{it}, as b_{kt} is not correlated with b_{it}, which was already shown before. Given representation of VAR is called a structural form. But due to the facts that, firstly, this form is not as easy to estimate and, secondly, it cannot be used for forecasting purposes, in time series analyses preference is usually given to the reduced form of VAR.

2.5 Pros and Cons of VAR Models

One should necessary understand basic pros and cons of VAR models, as they give a good insight on the specifics of the VAR modeling. The most significant benefits are:

- because of the fact that usually in the standard VAR model all the variables are endogenous, there is no sense in specifying, whether given variable is exogenous or endogenous;
- empirical studies show that VAR models are usually better for the prediction purposes than multiple simultaneous equations (MSR);
- parameters of the unrestricted reduced VAR form might be obtained using ordinary least squares method, which to some extent brings simplicity;

However, some problems are also connected with the use of VAR models:

- number of the parameters of the model might increase dramatically up to k^2, where k is the number of equations;
- VAR models are considered to be atheoretical, in a sense that they hardly explain the gist of the problem;
- VAR construction assumes that each of the AR processes is stationary, however transformations needed to achieve stationary sometimes may lead to the loss of information regarding long run relationships between time series.

- All of the economic variables highly depend on each other. However, there is no way to include all of them. Omitting important ones may lead to the specification error.

Creation of the VAR model according to Hušek (2007) can be represented in the following steps:

1. Ensure that all of the time series are stationary.
2. Determine the variables and the maximum lag length.
3. Possible restrictions on the parameters.
4. Residual tests and/or transformations of the vector of shocks.

First step will be discussed in a chapter on the stability, stationarity and the integration of the processes.

Second point will be a subject of the parts dedicated to the estimation and identification.

Possible restriction of the parameters is based on the fact that the number of the parameters to estimate is growing fast with the increase in the lags length; thus one might apply some restrictions in order to "save" extra degrees of freedoms. This will also be explained in more details in estimation and identification parts of the paper.

And finally, residual tests are needed to ensure that the estimated parameters get the properties of the best linear unbiased estimates (BLUEs).

2.6 Discussion on Forecasting

Necessary assumption for an acceptable forecast is that a tendency of a time series prevails in the future periods. More formally this statement can be represented as:

$$\hat{y}_{t+h} = f_1(y_t, y_{t-1}, y_{t-2}, \dots),$$

where $\hat{y}_{t+h}$ is a forecast for a period $t+h$, y_t is a value of a variable in a period t and $f(y_t, y_{t-1}, y_{t-2}, \dots)$ is a suitable function of the past values of y_t.

Often for a good forecast it is also necessary to take into account not only past values of the forecasted variable, but values of other variables too. Then the forecast function is

$$\hat{y}_{1,t+h} = f_1\big(y_{1,t}, y_{2,t}, \dots, y_{1,t-1}, y_{2,t-1}, \dots\big).$$

Similarly, a forecast function for a K-th variable is:

$$\hat{y}_{K,t+h} = f_K\big(y_{1,t}, y_{2,t}, \dots, y_{1,t-1}, y_{2,t-1}, \dots\big).$$

A set of time series y_{kt}, $k = 1 \dots K, t = 1 \dots T$ is called multiple time series. Important part of a multiple time series analysis is to determine acceptable $f_1, f_2, \dots, f_K$ functions, which are later on used to get forecasts.

Forecaster often needs to predict the future values of the vector $\boldsymbol{y_t}$. For this purposes the data generation process and information set Ω_t are needed. Ω_t contains all of the available information in the period t. Data generation process can be easily represented by VAR(p), information set can be thus formally represented by $\Omega_t = \{\boldsymbol{y_s} | s \leq t\}$. It should be once more noted, that the VAR forecast approach is atheoretical, meaning that there was no use of economic theory to specify relationships between the variables. Forecast origin is a period at which forecast will be made. The number of the periods for which forecast is made is usually called the forecast horizon.

The 1-step ahead forecast for a VAR(p) at the time of origin h is:

$$\boldsymbol{y_h}(1) = \boldsymbol{\varphi_0} + \sum_{\boldsymbol{i=1}}^{\boldsymbol{p}} \boldsymbol{\Phi_i y_{t+1-i}}$$

The corresponding forecast error is $\boldsymbol{e_h}(1) = \boldsymbol{\varepsilon_{t+1}}$ with a covariance matrix $\boldsymbol{\Sigma}$. The 2-step ahead forecast is represented by:

$$\boldsymbol{y_h}(2) = \boldsymbol{\varphi_0} + \sum_{\boldsymbol{i=2}}^{\boldsymbol{p}} \boldsymbol{\Phi_i y_{t+2-i}} + \boldsymbol{\Phi_1 y_h(1)}$$

With a forecast error of $\boldsymbol{e_h}(2) = \boldsymbol{\Phi_1 \varepsilon_{t+1}} + \boldsymbol{\varepsilon_{t+2}}$ and a covariance matrix of $\boldsymbol{\Sigma} + \boldsymbol{\Phi_1 \Sigma \Phi_1'}$. In a similar fashion one can continue to iterate. Note that, if y_t is weakly stationary, it eventually converges to its mean.

But due to the fact that this paper primarily deals with the structural analysis, there is no need to describe forecast technics any further.

2.7 Stability of Model

The stability of the model is studied through the reaction of the endogenous variables onto the exogenous shock. In the case if the process is stable, then shocks have a declining effect, which lasts for a relatively small amount of time, otherwise the process is not stable. If the process is integrated, then the effects of shocks never disappear and, finally, if the process is explosive, then shocks dramatically increase as the time passes.

Let us start with a simpler VAR(1):

$$\boldsymbol{y_t} = \boldsymbol{\varphi_0} + \boldsymbol{\Phi y_{t-1}} + \boldsymbol{\varepsilon_t} \,,$$

Later on all the conclusions would be extended to a VAR(p) model. It is necessary to understand the mechanism of data generation. Let's assume that the process begins at $t = 1$, we get:

$$\boldsymbol{y_1} = \boldsymbol{\varphi_0} + \boldsymbol{\Phi y_0} + \boldsymbol{\varepsilon_1}\ ,$$

$$\boldsymbol{y_2} = \boldsymbol{\varphi_0} + \boldsymbol{\Phi y_1} + \boldsymbol{\varepsilon_2} = \boldsymbol{\varphi_0} + \boldsymbol{\Phi}(\boldsymbol{\varphi_0} + \boldsymbol{\Phi y_0} + \boldsymbol{\varepsilon_1}) + \boldsymbol{\varepsilon_2} =$$

$$= (\boldsymbol{I} + \boldsymbol{\Phi})\boldsymbol{\varphi_0} + \boldsymbol{\Phi}^2\boldsymbol{y_0} + \boldsymbol{\Phi\varepsilon_1} + \boldsymbol{\varepsilon_2}\ ,$$

$$\dots$$

$$\boldsymbol{y_t} = (\boldsymbol{I} + \boldsymbol{\Phi} + \cdots + \boldsymbol{\Phi}^{t-1})\boldsymbol{\varphi_0} + \boldsymbol{\Phi}^t\boldsymbol{y_0} + \sum_{j=0}^{t-1}\boldsymbol{\Phi}^j\boldsymbol{\varepsilon_{t-j}}\ .$$

Which means that vectors $\boldsymbol{y_1}, \dots, \boldsymbol{y_t}$ are determined by $\boldsymbol{\varepsilon_1}, \dots, \boldsymbol{\varepsilon_t}$ and $\boldsymbol{y_0}$. But what if the process had started in the "infinite past"? After a slight modification we end up with:

$$\boldsymbol{y_t} = \boldsymbol{\varphi_0} + \boldsymbol{\Phi y_{t-1}} + \boldsymbol{\varepsilon_t} =$$
$$= (\boldsymbol{I} + \boldsymbol{\Phi} + \cdots + \boldsymbol{\Phi}^i)\boldsymbol{\varphi_0} + \boldsymbol{\Phi}^{i+1}\boldsymbol{y_{t-i-1}} + \sum_{i=0}^{\infty}\boldsymbol{\Phi}^i\boldsymbol{\varepsilon_{t-i}}\ .$$

Now it is worth mentioning that in the case, when all eigenvalues of $\boldsymbol{\Phi}$ are in absolute value less than 1,$(\boldsymbol{I} + \boldsymbol{\Phi} + \cdots + \boldsymbol{\Phi}^i)\boldsymbol{\varphi_0} \underset{i\to\infty}{\longrightarrow} \frac{\boldsymbol{\varphi_0}}{\boldsymbol{I}-\boldsymbol{\Phi}}$, also $\boldsymbol{\Phi}^{i+1}$ converges to zero, moreover $\sum_{i=0}^{\infty}\boldsymbol{\Phi}^i\boldsymbol{\varepsilon_{t-i}}$ exists in mean square, as *i* approaches infinity. So we can rewrite previous solution for $\boldsymbol{y_t}$ as

$$\boldsymbol{y_t} = \boldsymbol{\mu} + \sum_{i=0}^{\infty}\boldsymbol{\Phi}^i\boldsymbol{\varepsilon_{t-i}} = \boldsymbol{\mu} + \left(\sum_{i=0}^{\infty}\boldsymbol{\Phi}^i\boldsymbol{L}^i\right)\boldsymbol{\varepsilon_t}$$

Where $\boldsymbol{\mu} = \frac{\boldsymbol{\varphi_0}}{\boldsymbol{I}-\boldsymbol{\Phi}} = \begin{bmatrix}\bar{y}_1\\ \bar{y}_2\end{bmatrix}$

The distribution of $\boldsymbol{y_t}$ is uniquely determined by the distribution of $\boldsymbol{\varepsilon_t}$. The first and the second moments of $\boldsymbol{y_t}$ are:

$$E(\boldsymbol{y_t}) = (\boldsymbol{I} - \boldsymbol{\Phi})^{-1}\,\boldsymbol{\varphi_0} = \boldsymbol{\mu}$$
$$\Gamma_{\boldsymbol{y}}(h) = E(\boldsymbol{y_t} - \boldsymbol{\mu})(\boldsymbol{y_{t-h}} - \boldsymbol{\mu})' =$$
$$= \sum_{i=0}^{\infty}\boldsymbol{\Phi_{h+i}}\boldsymbol{\Sigma}\,(\boldsymbol{\Phi_i})'\ ,$$

Where $\boldsymbol{\Sigma} = E(\boldsymbol{\varepsilon_t}, \boldsymbol{\varepsilon_t'})$.

We call a VAR(1) process stable if all eigenvalues of $\boldsymbol{\Phi}$ are in absolute value lower than 1. It can also be shown that the same conclusions about stability can be made under the equal assumption that $det(\boldsymbol{I} - \boldsymbol{\Phi}\mathrm{z}) = det\left(I - \Phi_1\mathrm{z} - \cdots - \Phi_p\mathrm{z}^p\right) \neq 0$ for $|z| \leq 1$. This polynomial is called reversed characteristic polynomial of the VAR(p). If the roots of the reversed characteristic equation lie strictly outside the unit circle, VAR(p) is considered to be stable.

It is true that if stability condition is satisfied, given VAR(1) process is stationary. All of the previous properties of a VAR(1) can be easily extended to VAR(p), as any VAR(p) can be written in the form of VAR(1). It will be shown later.

Another view on the stability condition consists in the use of lag operators.

$$y_{1t} = \varphi_1 + \Phi_{11} L y_1 + \Phi_{12} L y_2 + \varepsilon_{1t},$$
$$y_{2t} = \varphi_2 + \Phi_{21} L y_1 + \Phi_{22} L y_2 + \varepsilon_{2t},$$

Or equally:

$$(1 - \Phi_{11} L) y_{1t} = \varphi_1 + \Phi_{12} L y_2 + \varepsilon_{1t},$$
$$(1 - \Phi_{22} L) y_{2t} = \varphi_2 + \Phi_{21} L y_1 + \varepsilon_{2t},$$

From the last equation $L y_{2t}$ is:

$$L y_{2t} = L \frac{(\varphi_2 + \Phi_{21} L y_1 + \varepsilon_{2t})}{(1 - \Phi_{22} L)}$$

Substituting it into the $(1 - \Phi_{11} L) y_{1t} = \varphi_1 + \Phi_{12} L y_2 + \varepsilon_{1t}$ equation, we get:

$$(1 - \Phi_{11} L) y_{1t} = \varphi_1 + \Phi_{12} L \frac{(\varphi_2 + \Phi_{21} L y_1 + \varepsilon_{2t})}{(1 - \Phi_{22} L)} + \varepsilon_{1t}$$

VAR(1) has been transformed into second order difference equation**.** Solving it for y_{1t} :

$$y_{1t} = \frac{\Phi_{10}(1 - \Phi_{22} L) + \Phi_{12}\Phi_{20} + (1 - \Phi_{22} L)\varepsilon_{1t} + \Phi_{12}\varepsilon_{2t-1}}{(1 - \Phi_{11} L)(1 - \Phi_{22} L) - \Phi_{12}\Phi_{21} L^2}$$

Solution for y_{2t} can be obtained in a similar way. Both of y_{1t} and y_{2t} have the same characteristic equations. For the convergence, roots of inversed characteristic equation $(1 - \Phi_{11} L)(1 - \Phi_{22} L) - \Phi_{12}\Phi_{21} L^2$ should lie outside the unit circle. If Φ_{12} and Φ_{21} do not equal zero, the characteristic roots will also be the same, hence both y_{1t} and y_{2t} will have a similar time path.

2.8 Moving Average Representation of VAR Process

Wold decomposition theorem, named after Herman Wold, proves that any zero-mean process can be uniquely represented as the sum of a stochastic process in a form of infinite moving average and a linearly predictable deterministic process.

VAR can be represented in the form of vector moving average (VMA). It can trace the time path of the effect of shocks (innovations) on the variables. Alternatively, y_{1t} and y_{2t} are expressed by the current and past values of ε_{1t} and ε_{1t}. In fact, we have already seen VMA in the previous discussion on the topic of stability. If VAR satisfies the stability condition, then it can be represented as following:

$$\boldsymbol{y_t} = \boldsymbol{\mu} + \sum\nolimits_{\boldsymbol{i=0}}^{\infty} \boldsymbol{\Phi^i \varepsilon_{t-i}}$$

Coefficients of $\boldsymbol{\Phi^i}$ may be considered as impact multipliers, as they pass the effect of the shock onto $\boldsymbol{y_t}$. For example, consider a shock ε_{1t} at a time t. It will be fully passed onto y_{1t}, as $\boldsymbol{\Phi}^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. At $t+1$ the effect will be incomplete, as only $\Phi_{11}\varepsilon_{1t}$ will be passed etc. It is possible to construct impulse response functions in a similar manner. Situation will be slightly different, if the variables would be allowed to have contemporaneous effect on each other.

Alternatively, VMA form is:

$$(\boldsymbol{y_t} - \boldsymbol{\mu}) = \boldsymbol{\varepsilon_t} + \boldsymbol{\Phi\varepsilon_{t-1}} + \boldsymbol{\Phi^2\varepsilon_{t-2}} + \boldsymbol{\Phi^3\varepsilon_{t-3}} + \cdots$$

As $\boldsymbol{\varepsilon_t}$ is serially uncorrelated, then $Cov(\boldsymbol{\varepsilon_t}, \boldsymbol{y_t}) = \boldsymbol{0}$. Moreover, $\boldsymbol{\varepsilon_t}$ is not correlated with $\boldsymbol{y_{t-l}}$ for $l > 0$. That is why ε_t is called shock. Now let me determine the autocovariances of $\boldsymbol{y_t}$ from the VMA representation:

$$Cov(\boldsymbol{y_t}) = \Gamma_y(0) = \boldsymbol{\Sigma} + \boldsymbol{\Phi\Sigma\Phi'} + \boldsymbol{\Phi^2\Sigma}(\boldsymbol{\Phi^2})' + \cdots = \sum_{\boldsymbol{i=0}}^{\infty} \boldsymbol{\Phi^i\Sigma}(\boldsymbol{\Phi^i})$$

From VMA we can obtain MA representation for y_t. We just need to multiply the last equation by $k \times k$ matrix $J = [I: 0: \dots: 0]$:

$$\begin{aligned} y_t &= \boldsymbol{Jy_t} = \boldsymbol{J\mu} + \sum\nolimits_{\boldsymbol{i=0}}^{\infty} \boldsymbol{J\Phi^i J' J \varepsilon_{t-i}} \\ &= \mu + \sum\nolimits_{\boldsymbol{i=0}}^{\infty} \Phi_i \varepsilon_{t-i} \end{aligned}$$

Where $\mu = \boldsymbol{J\mu}$, $\Phi_i = \boldsymbol{J\Phi^i J'}$, $\varepsilon_{t-i} = \boldsymbol{J\varepsilon_{t-i}}$. The last representation is often called canonical error representation.

Coefficients of canonical error representation can be computed easily applying lag operators over corresponding VAR model.

VMA representation of VAR(1) process in the matrix form is:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \varphi_{10} \\ \varphi_{20} \end{bmatrix} + \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix} + \sum_{i=0}^{\bowtie} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}^i \begin{bmatrix} \varepsilon_{1,t-i} \\ \varepsilon_{2,t-i} \end{bmatrix}$$

It is also possible to define moving average representation of the VAR(p) model using lag operators:

$$\left(\boldsymbol{I} - \Phi_{\boldsymbol{1}} L - \Phi_{\boldsymbol{2}} L^2 - \cdots - \Phi_{\boldsymbol{p}} L^p\right) y_t = \boldsymbol{\varphi_0} + \boldsymbol{\varepsilon_t}$$
$$\boldsymbol{y_t} = \left(\boldsymbol{I} - \Phi_{\boldsymbol{1}} L - \Phi_{\boldsymbol{2}} L^2 - \cdots - \Phi_{\boldsymbol{p}} L^p\right)^{-1} (\boldsymbol{\varphi_0} + \boldsymbol{\varepsilon_t})$$
$$\boldsymbol{y_t} = \left(\sum_{i=0}^{\infty} \boldsymbol{A_i} L^i\right)(\boldsymbol{\varphi_0} + \boldsymbol{\varepsilon_t}) = \left(\sum_{i=0}^{\infty} \boldsymbol{A_i} L^i\right)\varphi_0 + \left(\sum_{i=0}^{\infty} \boldsymbol{A_i} L^i\right)\varepsilon_t$$

where $\boldsymbol{A_0} = 1$, $\boldsymbol{A_i}$ are constant matrices $k \times k$, so that $\boldsymbol{A}(L)\Phi(L) = \boldsymbol{I}$.

2.9 Stationarity

A process is strictly stationary if its properties are unaffected by a change of time. Analogously, the joint probability distribution is said to stay the same regardless time changes. A process is weakly stationary (or covariance stationary) in the case that its first and second moment are time invariant and finite (no longer the entire distribution is supposed to be time-invariant, as in the case of strict stationarity). Formally,

1. $E(\boldsymbol{y_t}) = \boldsymbol{\mu} < \infty$ for all t
2. $E(\boldsymbol{y_t} - \boldsymbol{\mu})(\boldsymbol{y_{t-h}} - \boldsymbol{\mu})' = \Gamma(h) = \Gamma(-h)'$

In other words, all of the y_t have the same finite mean vector for all *t* and the autocovariances are not supposed to be depended on *t*. A process is considered to be strictly stationary if all finite-dimensional distributions are time-invariant. Weakly asymptotically stationary process is a process that starts at a time of origin and which expectations of $\boldsymbol{y_t}$ and autocovariances converge to finite limits.

As already have been mentioned, stability implies stationarity, not vice versa.

2.10 Equivalence of VAR(1) and VAR(p)

VAR(p) can be represented in the form of VAR(1). This is a tremendously important way of simplification of the model.

For example, consider the following VAR(p) process:

$$\boldsymbol{y_t} = \left(\boldsymbol{\Phi_1} L + \boldsymbol{\Phi_2} L^2 + \cdots + \boldsymbol{\Phi_p} L^p\right)\boldsymbol{y_t} + \boldsymbol{\varphi_0} + \boldsymbol{\varepsilon_t}$$

It can be transformed into a kp-dimensional VAR(1) model:

$$\boldsymbol{y_t} = \boldsymbol{\Phi}^* \boldsymbol{y_{t-1}} + \boldsymbol{S_t} + \boldsymbol{W_t}$$

Where

$$\boldsymbol{y_t} = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}, \boldsymbol{\Phi}^* = \begin{bmatrix} \boldsymbol{\Phi_1} & \boldsymbol{\Phi_2} & & \boldsymbol{\Phi_{p-1}} & \boldsymbol{\Phi_p} \\ \boldsymbol{I_n} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I_n} & & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & & \ddots & & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{I_n} & \boldsymbol{0} \end{bmatrix}, \boldsymbol{S_t} = \begin{bmatrix} \boldsymbol{\varphi_0} \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix}, \boldsymbol{W_t} = \begin{bmatrix} \boldsymbol{\varepsilon_t} \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix}$$

$\boldsymbol{y_t}$, $\boldsymbol{S_t}$, $\boldsymbol{W_t}$ are $kp \times 1$ vectors and $\boldsymbol{\Phi}^*$ is a $kp \times kp$ matrix. $\boldsymbol{\Phi}^*$ matrix is usually called the companion matrix. Please note, that the both reversed characteristic equations of an original VAR(p) and modified VAR(1) are the same.

2.11 Integrated Process

A Process is called integrated of order one, if its first differences form a stationary process. A process is integrated of order *n* if:

$$(I - L)^n y_t = y_t{}'$$

Where y_t' is a stationary process.

Characteristic equation of an integrated process of order *n* has *n* roots equal to 1. Please note that shocks in the integrated processes have a permanent effect, as it can be shown that the integrated process can be factorized as the sum of deterministic trend, stochastic trend and a cyclic stationary process.

The main difference between processes that are said to be integrated of order 0 and order 1 is that $I(0)$ series is mean reverting (it is expected to return to its mean in a long run), whereas $I(1)$ fluctuates widely. $I(0)$ is said to have limited memory on the past behavior, $I(1)$ in its turn has an infinite long memory, which means that a single innovation has a permanent effect on the process. This idea is clearly seen in the autocorrelation function: in the case of $I(0)$ autocorrelations decline, as lag increases, concerning $I(1)$ autocorrelations decline to zero very slowly.

Also one needs to check, whether $I(1)$ variables does not have a cointegrational relationship, which is I(0). As follows, it will not be appropriate to use first differentials and not only will it lead to a mistake in a specification, but also the resulting model will not be capable of providing any information regarding long-term relationships of the given variables.

To test for a unit root in a time series sample Dickey-Fuller statistic might be applied (or also it's augmented version). Dickey-Fuller (DF) test tests, whether investigated time series is $I(1)$ under the null hypothesis against the alternative that time series is $I(0)$. The more negative value test statistic gets, the lower is the chance of not rejecting the null hypothesis that there is a unit root.

Alternatively, Phillips-Perron (PP) unit root test with the same null hypothesis that there is a unit root existence might be used. In a sense, PP statistic is a modification of DF statistic, which is robust to serial correlation in errors.

As a critique for these tests, it should be noted that they have very low power against $I(0)$ alternatives which are being close to $I(1)$.

Stationarity tests, in contrast to DF and PP, are for the null hypothesis that given time series is $I(0)$. KPSS test (named after Kwiatkowski, Phillips, Schmidt and Shin) is especially popular nowadays. One should always consider an option of making a decision about unit roots only after addressing to all of the tests.

2.12 Autocovariances and Autocorrelations of Stable VAR Process

In practice it is often not very convenient to use the formula discussed earlier:

$$\Gamma_y(h) = \sum_{\boldsymbol{i=0}}^{\infty} \boldsymbol{\Phi}_{\boldsymbol{h+i}} \boldsymbol{\Sigma} \, (\boldsymbol{\Phi}_{\boldsymbol{i}})'$$

Fortunately, VAR coefficient matrices provide a way to compute autocovariances. Consider a following VAR process:

$$\boldsymbol{y_t} = \boldsymbol{\varphi_0} + \boldsymbol{\Phi} \boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t} \, ,$$

If $(\boldsymbol{I} - \boldsymbol{\Phi})$ matrix is regular, then it is possible to rewrite VAR process in a mean adjusted form:

$$\boldsymbol{y_t} - \boldsymbol{\mu} = \boldsymbol{\Phi}(\boldsymbol{y_{t-1}} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon_t}$$

Multiplying both sides of the equation by the $(\boldsymbol{y_{t-h}} - \boldsymbol{\mu})'$ and then taking the expectation, we get:

$$E[(\boldsymbol{y_t} - \boldsymbol{\mu})(\boldsymbol{y_{t-h}} - \boldsymbol{\mu})'] = \boldsymbol{\Phi}\mathrm{E}[(\boldsymbol{y_{t-1}} - \boldsymbol{\mu})(\boldsymbol{y_{t-h}} - \boldsymbol{\mu})'] + \boldsymbol{E}[\boldsymbol{\varepsilon_t}(\boldsymbol{y_{t-h}} - \boldsymbol{\mu})']$$

Where $\mathrm{E}[\boldsymbol{\Phi}(\boldsymbol{y_{t-1}} - \boldsymbol{\mu})(\boldsymbol{y_{t-h}} - \boldsymbol{\mu})'] = \boldsymbol{\Phi}\mathrm{E}[(\boldsymbol{y_{t-1}} - \boldsymbol{\mu})(\boldsymbol{y_{t-h}} - \boldsymbol{\mu})']$, as $\boldsymbol{\Phi}$ is a matrix of known coefficients. For $h = 0$:

$$\Gamma_{\boldsymbol{y}}(0) = \boldsymbol{\Phi}\Gamma_y(-1) + \Sigma = \boldsymbol{\Phi}\Gamma_y(1)' + \Sigma$$

For $h > 0$:

$$\Gamma_{\boldsymbol{y}}(h) = \boldsymbol{\Phi}\Gamma_{\boldsymbol{y}}(h-1)$$

Or also:

$$\Gamma_{\boldsymbol{y}}(h) = \boldsymbol{\Phi}^{\boldsymbol{h}}\Gamma_{\boldsymbol{y}}(0)$$

These equations are known as Yule-Walker equations. Their implication consists in the fact that if $\boldsymbol{\Phi}$ and $\Sigma = \Gamma_{\boldsymbol{y}}(0)$ are known, then it is possible to calculate $\Gamma_{\boldsymbol{y}}(h)$**.**

Using The Kronecker product and the vec operator, it is convenient to determine $\Gamma_{\boldsymbol{y}}(0)$, if $\boldsymbol{\Phi}$ and $\Sigma = \Gamma_{\boldsymbol{y}}(0)$ are given. As $\Gamma_{\boldsymbol{y}}(1) = \boldsymbol{\Phi}\Gamma_{\boldsymbol{y}}(0)$, we get:

$$\Gamma_{\boldsymbol{y}}(0) = \boldsymbol{\Phi}\Gamma_{\boldsymbol{y}}(0)\boldsymbol{\Phi}' + \boldsymbol{\Sigma}$$

Analogically:

$$\begin{aligned} vec\Gamma_{\boldsymbol{y}}(0) &= vec\left(\boldsymbol{\Phi}\Gamma_{\boldsymbol{y}}(0)\boldsymbol{\Phi}'\right) + vec\boldsymbol{\Sigma} \\ &= (\boldsymbol{\Phi} \otimes \boldsymbol{\Phi})vec\left(\Gamma_{\boldsymbol{y}}(0)\right) + vec\boldsymbol{\Sigma} \end{aligned}$$

Continuing:

$$vec\Gamma_{\boldsymbol{y}}(0) = (\boldsymbol{I}_{\boldsymbol{k}^2} - \boldsymbol{\Phi} \otimes \boldsymbol{\Phi})^{-\mathbf{1}}vec\boldsymbol{\Sigma}$$

It can be shown that, if the $\boldsymbol{y_t}$ is stable, then $\boldsymbol{I}_{\boldsymbol{k}^2} - \boldsymbol{\Phi} \otimes \boldsymbol{\Phi}$ is invertible. When we get $\Gamma_{\boldsymbol{y}}(0)$, we might use $\Gamma_{\boldsymbol{y}}(h) = \boldsymbol{\Phi}\Gamma_{\boldsymbol{y}}(\boldsymbol{h}-\mathbf{1})$ to calculate $\Gamma_{\boldsymbol{y}}(h)$ for $h = 1,2,3$ ... .

As autocorrelations are in a sense independent on the unit of measure versions of autocovariances, they are usually more preferred to work with. Autocorrelations are defined as:

$$\boldsymbol{R}_y(h) = \left[r_{ij}(0)\right] = \boldsymbol{\rho_0} = [\rho_{i,j}(0)] = D^{-1}\Gamma_{\boldsymbol{y}}(0)D^{-1}$$

Where D is a diagonal matrix, so that diagonal elements are square roots of the diagonal elements of $\Gamma_{\boldsymbol{y}}(0)$, in other words, standard deviations (std) of $\boldsymbol{y_t}$. The i,jth element of the concurrent (concurrent stands for the fact, that it demonstrates the correlations between time series at time t) correlation matrix or simply the correlation coefficient between y_{it} and y_{jt} is:

$$\rho_{i,j}(0) = \frac{Cov(y_{it}, y_{jt})}{std(y_{it})std(y_{jt})}$$

Correlation matrix is symmetric with unit diagonal elements.

2.13 Vectorization and Kronecker Product

As it might be easier to calculate autocovariances with a help of the vectorization and the Kronecker product, a part of this paper will be devoted to the basic explanation of these terms.

Denote $\boldsymbol{A}$ as a $m \times n$ matrix and $\boldsymbol{B}$ as a $x \times z$ matrix, then the Kronecker product of $\boldsymbol{A}$ and $\boldsymbol{B}$ is $mx \times nz$ matrix:

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{bmatrix}$$

Where $\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$

The $vec\ operator$ applied over the $\boldsymbol{A}$ matrix transforms it into a $mn \times 1$ vector. It basically stacks the columns of the $\boldsymbol{A}$ matrix, so that:

$$vec(A) = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m-1,n} \\ a_{mn} \end{bmatrix}$$

2.14 Estimation and Identification

The variables should be selected according to the economic theory. For determination of the appropriate lag length one should use lag-length tests (would be discussed later). $k + pk^2$ coefficients are to be estimated (where k is a number of variables and p is a number of lags). A lot of lags fastly consume degrees of freedom. No doubt, that VAR process might be overparameterized, so that it is possible to a priori impose the maximum lag length, despite of the test or other suggestions. However, this, on the other hand, may lead to a loss of relevant information.

It is important to note, that t-tests on individual coefficients may not be reliable, as regressors are likely to be collinear, it would be better to use F test for the whole model. As already been noted, because there are only predetermined variables on the right side of equation and error terms are supposed to be uncorrelated, one can simply use OLS for the estimation of the parameters. OLS estimates have important properties of consistency and asymptotic efficiency. If some of equations have regressors not included in others, one should use seemingly unrelated regressors to get efficient estimates of the coefficients. Maximum likelihood method can also be used. These two methods are asymptotically equivalent.

Lots of discussions have been made on the topic, whether it is necessary for the variables to be stationary. Some claim that one should not use differencing even if the variable contain unit root, as it destroys comovements in the data.

Also, it is recommended not to use detrending, as the majority view is that VAR should be as similar to the original data generating process as possible.

In order to check, that there is no correlation among residual series, it is useful to address to the $Q_k(m)$ statistic. $Q_k(m)$ statistic asymptotically has a χ^2 distribution with $k^2m - g$ degrees of freedom, where g represents the number of the estimated parameters.

2.15 Estimation of the Number of Lags

Generally, increase in the order of the model, reduces the size of the residuals, but on the other hand also decreases forecasting ability of the model.

It is worth mentioning, that we assume that the original data process is generated by the VAR mechanism.

The residual covariance matrix $\boldsymbol{\Sigma}$ of VAR(p) can be estimated as:

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{T - 2p - 1} \sum_{t=p+1}^{T} \widehat{\boldsymbol{\varepsilon}}_t (\widehat{\boldsymbol{\varepsilon}}_t)'$$

To choose the length of lags, one can test the hypothesis $H_0 \colon \boldsymbol{\Phi_p} = 0$ against $H_1 \colon non\ H_0$. The corresponding test statistic is

$$M(p) = -(T - k - p - \frac{3}{2}) \ln \left(\frac{|\widehat{\boldsymbol{\Sigma}_p}|}{|\widehat{\boldsymbol{\Sigma}}_{p-1}|} \right)$$

Where $|\widehat{\boldsymbol{\Sigma}}_p|$ is a determinant of the residual covariance matrix, p is an order of the model.

In general, VAR(p) is being tested against VAR($p - 1$). This statistic asymptotically has a χ^2 distribution with k^2 degrees of freedom.

Alternatively, one should begin with a longest possible lag, keeping in mind degrees of freedom. Estimate the model, get variance/covariance matrix of the residuals. Later, repeat the mentioned steps with lower lags. Then, in order to compare two models, it is useful to address to likelihood ratio test.

The likelihood ratio statistic is:

$$(T)(\log|\Sigma_r| - \log|\Sigma_u|)$$

Where $|\Sigma_r|$ is again a determinant of the variance/covariance matrix of the residuals with less lags (restricted model) and $|\Sigma_u|$ is a determinant of variance/covariance matrix of the model with maximum lags (unrestricted model) and T is a number of observations. If the equations of the unrestricted model contain different number of regressors, it is useful to modify given statistic in a following way:

$$(T - c)(\log|\Sigma_r| - \log|\Sigma_u|)$$

Where c is a maximum number of regressors in the longest equation.

Given statistic has a χ^2 distribution with degrees of freedom equal to the number of the restrictions in the entire system (the restricted sum of lags through all of the equations). If this statistic has a significant value, it indicates that only lower number of lags is a binding restriction. If the value of statistic is less than related χ^2 at a chosen significance level, it is not possible to reject the null hypothesis of only lower number of lags.

Likelihood ratio test may not be very informative in the case of the small samples. Also it should be always used when comparing restricted version of the model with its unrestricted "version".

Thus, it may be instead useful to address to the multivariate generalizations of the Akaike information criterion (AIC), Bayesian information criterion (BIC) and Hannan-Quinn information criterion (HQIC). In this case $\boldsymbol{\Sigma}$ is estimated differently, according to Tsay (2005):

$$\breve{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}(\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{t}})'$$

Akaike information criterion (AIC) under the normality condition of the $\boldsymbol{\varepsilon_t}$ is defined as:

$$AIC(p) = \ln\left(\left|\breve{\boldsymbol{\Sigma}}_{\boldsymbol{p}}\right|\right) + \frac{2k^2p}{T}$$

Similarly, BIC (or also "SBC, SBIC, as it was developed by G. E. Schwarz, who gave Bayesian argument for adopting it" (wikipedia)) is defined as

$$BIC(p) = \ln\left(\left|\breve{\boldsymbol{\Sigma}}_{\boldsymbol{p}}\right|\right) + \frac{\ln(T)\, k^2p}{T}$$

And HQIC is

$$HQIC(p) = \ln\left(\left|\breve{\boldsymbol{\Sigma}}_{\boldsymbol{p}}\right|\right) + \frac{2\ln(\ln T)\, k^2 p}{T}$$

Optimal lag-length minimalizes those criterions. Note, that AIC usually suggests applying higher lag order, whereas BIC lower. All of the information criteria penalize the increase in the order of the VAR model.

2.16 Granger Causality

Concept developed by Granger can be, under suitable conditions, used in the context of VARs. Moreover, nowadays it is really popular. The idea behind causality is that cause cannot come after effect. Thus, cause can help to improve prediction of the effect. For example, if y_{1t} does not improve the forecasting of y_{2t}, then y_{1t} does not Granger cause y_{2t}.

In order to describe Granger causality mathematically suppose that Ω_t is a set of all the relevant information, $y_{2t}(h|\Omega_t)$ is a predictor of the process y_{2t} and $\Sigma_z(h|\Omega_t)$ is a MSE forecast. y_{1t}Granger causes y_{2t}, if:

$$\Sigma_z(h|\Omega_t) < \Sigma_z(h|\Omega_t \backslash \{y_{1s}|s \leq t\})$$

This formula basically repeats the previous discussion, as it states that y_{2t} can be predicted in a better way, when y_{1t} is taken into account. There also is a concept of the instant causality. It suggests that by including at a time t $y_{1,t+1}$ heplps to improve the forecast of $y_{2,t+1}$. It can be shown, that this concept is symmetric (it holds vise versa).

To test Granger causality it is possible to apply standard F-test:

$$\Phi_{21}(1) = \Phi_{21}(2) = \Phi_{21}(3) = \cdots = 0$$

In the case of k variable model, if one variable does not Granger cause another, then all of the regarding coefficients can be set equal to zero.

Another way of testing Granger causality is to address to a Wald test. The main idea behind it is similar. Eviews provides convenient ways of testing Granger causality. Granger causality is considered to be a “weaker term”, comparing it to the exogenity.

2.17 Impulse Response Analysis

The main subject of the study of impulse response analysis is a way that one variable responds to the shock (or impulse) of another variable in a higher dimensional system. Sometimes this analysis is referred to as a multiplier analysis. Consider uncorrelated shocks.

For example, it can be shown that a unit shock on the i-th variable after h-periods results in the effect on the other variables represented by the i-th column of $\boldsymbol{\Phi}^h$, all other thing being constant. That is why the elements of $\boldsymbol{\Phi}$ matrix are often called impulse responses or also dynamic (impact) multipliers. It may be convenient to plot such effects of the unit shock of one variable on another. Of course, due to the fact that the system's equations are being stable, such unit effects converge to zero quite rapidly.

Please note, that if one variable does not granger cause other variables (viewed as a set), innovation in such a variable does not have any effect on the other variables.

Surprisingly, if VAR(p) is a k-dimensional model and if "first $pk - p$ responses of variable j to an impulse in variable k are zero, all the following responses must also be zero" (Lütkepohl, 2005).

For simplicity of the notation, denote $\boldsymbol{\Phi}^h = \boldsymbol{A}_h$. If an object of interest is to calculate the accumulated responses over predetermined amount of periods (say, n) to the shock in another variable, one can use $\psi_n = \sum_{h=0}^{n} \boldsymbol{A}_h$.

Such an analysis may become problematic if the shocks in different variables are no longer assumed to be independent. Thus, it is no longer possible to quantify effect of the shock in a given variable on other variables, as such a shock may be accompanied with the shock in another variable in the same period due to correlation between them.

One then might want to use modified VMA representation, with the uncorrelated residuals. For this purposes it is useful to address to Cholesky decomposition (see the beginning of the paper) $\boldsymbol{\Sigma} = \boldsymbol{LGL}'$, where $\boldsymbol{L}$ is a lower triangular matrix with unit diagonal elements and $\boldsymbol{G}$ is a diagonal matrix.

If we denote $\boldsymbol{L} = \boldsymbol{PD}^{-\mathbf{1}}$ (or equally $\boldsymbol{P} = \boldsymbol{LD}$) and $\boldsymbol{G} = \boldsymbol{DD}'$, where $\boldsymbol{D}$ is a diagonal matrix, then $\boldsymbol{\Sigma} = \boldsymbol{PP}'$ holds. No more than simple matrix algebra is needed to prove this:

$$\boldsymbol{\Sigma} = \boldsymbol{PP}' = \boldsymbol{LD}(\boldsymbol{LD})' = \boldsymbol{LDD}'\boldsymbol{L}' = \boldsymbol{LGL}'$$

Consider zero mean VAR(p) process:

$$\boldsymbol{y_t} = \boldsymbol{\Phi_1 y_{t-1}} + \cdots + \boldsymbol{\Phi_p y_{t-p}} + \boldsymbol{\varepsilon_t}$$

Multiplying it by $\boldsymbol{\Lambda} = \mathbf{L}^{-\mathbf{1}}$ from the left, we get:

$$\boldsymbol{\Lambda y_t} = \boldsymbol{\Phi_1}^* \boldsymbol{y_{t-1}} + \cdots + \boldsymbol{\Phi_p}^* \boldsymbol{y_{t-p}} + \boldsymbol{v_t}$$

Where $\boldsymbol{\Phi_i}^* = \boldsymbol{\Lambda\Phi_1}$ and $\boldsymbol{\varepsilon_t} = \boldsymbol{\Lambda v_t}$ with a diagonal covariance matrix:

$$\boldsymbol{G} = E(\boldsymbol{v_t v_t'}) = \boldsymbol{\Lambda} E(\boldsymbol{\varepsilon_t \varepsilon_t}')\boldsymbol{\Lambda}' = \boldsymbol{\Lambda\Sigma\Lambda}'$$

After adding $(\boldsymbol{I}-\boldsymbol{\Lambda})\boldsymbol{y_t}$ to the both sides of equation, we get:

$$\boldsymbol{\Lambda y_t} + \boldsymbol{y_t} - \boldsymbol{\Lambda y_t} = (\boldsymbol{I}-\boldsymbol{\Lambda})\boldsymbol{y_t} + \boldsymbol{\Phi_1}^{*}\boldsymbol{y_{t-1}} + \cdots + \boldsymbol{\Phi_p}^{*}\boldsymbol{y_{t-p}} + \boldsymbol{v_t}$$

$$\boldsymbol{y_t} = \boldsymbol{\Phi_0}^{*}\boldsymbol{y_t} + \boldsymbol{\Phi_1}^{*}\boldsymbol{y_{t-1}} + \cdots + \boldsymbol{\Phi_p}^{*}\boldsymbol{y_{t-p}} + \boldsymbol{v_t}$$

As $\boldsymbol{L}$ is a lower triangular matrix with unit diagonal elements, then $\boldsymbol{\Lambda}$ will also have the same properties and it follows that:

$$\boldsymbol{\Phi_0}^{*} = \boldsymbol{I}-\boldsymbol{\Lambda} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \beta_{21} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_{k-1,1} & \beta_{k-1,2} & \cdots & 0 & 0 \\ \beta_{k1} & \beta_{k2} & \cdots & \beta_{k,k-1} & 0 \end{bmatrix}$$

Multiplication of $\boldsymbol{\Phi_1}^{*}$ by $\boldsymbol{y_t}$ shows that there is no instantaneous y_{1t} on the right side of the first equation. In such system there is no instantaneous effect of y_{st} on y_{kt}, if $k < s$.

This representation of VAR(p) is called a recursive model. Please note that in the case of the recursive models one has to specify the “instantaneous causal ordering of the variables. This type of causality is therefore sometimes referred to as Wold-causality”.

2.18 Variance Decomposition

To forecast a value of $\boldsymbol{y_{t+1}}$, given that the $\boldsymbol{\varphi_0}$, $\boldsymbol{y_t}$ and $\boldsymbol{\Phi}$ are known, we need to take the conditional expectation of y_{t+1}:

$$\boldsymbol{y_h}(1) = E(\boldsymbol{y_{t+1}}) = \boldsymbol{\varphi_0} + \boldsymbol{\Phi y_t}$$

This forecast will have an error of:

$$\boldsymbol{y_{t+1}} - E(\boldsymbol{y_{t+1}}) = \boldsymbol{\varepsilon_t}$$

Analogously, n-step ahead forecast can be represented by:

$$\boldsymbol{y_h}(n) = E(\boldsymbol{y_{t+n}}) = (\boldsymbol{I} + \boldsymbol{\Phi} + \boldsymbol{\Phi^2} + \cdots + \boldsymbol{\Phi^{n-1}})\boldsymbol{\varphi_0} + \boldsymbol{\Phi^n y_t}$$

Its forecast error is:

$$\boldsymbol{\varepsilon_{t+n}} + \boldsymbol{\Phi\varepsilon_{t+n-1}} + \boldsymbol{\Phi^2\varepsilon_{t+n-2}} + \cdots + \boldsymbol{\Phi^{n-1}\varepsilon_{t+1}}$$

If we focus solely on y_{1t}, it’s variance of n-step ahead forecast error ($\sigma_{y_1}(n)^2$) is

$$\sigma_{y_1}(n)^2 = \sigma_{y_1}^2(\Phi_{11}(0)^2 + \Phi_{11}(1)^2 + \cdots + \Phi_{11}(n-1)^2) + \\ + \sigma_{y_2}^2(\Phi_{12}(0)^2 + \Phi_{12}(1)^2 + \cdots + \Phi_{12}(n-1)^2)$$

Variance of forecast error increases, as we consider longer forecast horizon. Note, that it is possible to decompose the proportions of $\sigma_{y_1}(n)^2$ due to shocks in both of y_{1t} and y_{2t}.

Variance decomposition describes the proportion of the movements in a sequence due to corresponding shock versus shock of the other variable. It is possible to state that y_{1t} sequence is exogenous, if ε_{2t} shocks do not explain forecast error variance of y_{1t}. Thus y_{1t} evolves independently of ε_{2t} shocks of the other variable.

3 Exchange Rate Pass-Through

Exchange rate pass-through (ERPT) can be defined as the change in the domestic prices (whether import, consumer or wholesaler prices, measured by import price index (IPI), consumer price index (CPI) or wholesale price index (WPI)), originally caused by the volatility of the exchange rate. Exchange rate fluctuations affect economic activity in several ways. Firstly, “import prices transmit an exchange rate shock into domestic inflation directly, via imported goods, which constitute a part of final consumption” (Babecká-Kucharčuková, 2009). And, secondly, indirectly, via imported semiproducts for goods produced domestically.

There is no doubt, that ERPT plays an incredibly important role in the macroeconomic policy design, as it deals with different types of inflation, which means that it should be one of the most important concerns of national banks. As in the year of 1998 the Czech National Bank (ČNB) has switched to an inflation targeting, exchange rates are to be followed, as they effect domestic inflation and thus may spoil ČNB’s predictions.

Also, due to an increased volatility of exchange rates after introducing floating exchange rates and closer relationships between countries as a result of globalization it is better to know if prices will also have a high volatility. Moreover, ERPT is a big concern regarding small open economies, particularly Czech Republic, as they might be affected the most.

For illustrative purposes, imagine that some country’s currency depreciates. Theoretically this is supposed to be one of the mechanisms to support export, thus also to improve external balance. Although, if domestic prices react to the change of the nominal exchange rate in the comparable manner, exporters might not get any competitive advantage over foreign companies. They can possibly even get worse in case when they hold vast foreign currency liabilities. The same principle might be applied vise versa. One should expect a lower export, as a result of the appreciation of a currency. But if domestic prices “adjust” on time nothing will happen, all other things being constant. Hence, one needs to clearly understand how does volatility of exchange rates affects different types of prices to analyze possible changes in the trade balance.

Exchange rate is generally determined by a large number of different parameters. Starting from inflation, interest rate and ending with the business climate or investment opportunities. It is safe to state, that almost every economic variable might somehow affect nominal exchange rate. But let’s focus on the most important factors that might determine exchange rate.

Lots of studies have shown, that higher inflation mostly depreciates the currency, as people (or investors, companies) usually try to keep their savings in different currencies due to high inflation expectations, by that they raise the demand for other currencies and lower the demand for the domestic currency. This is supposed to result in the depreciation of the domestic currency.

It is generally believed that changes in the interest rate also have a significant influence on the exchange rate, as they might whether attract new investors from abroad, thus increase a demand for domestic currency or influence behavior of firms.

However, some studies show that exchange rate is exogenous to a large number of macroeconomic variables.

Complete ERPT assumes that an x-percent depreciation of a currency is fully passed, which will result in an x-percent increase in that currency price of the imported good. In the case of an incomplete ERPT increase in the price might be smaller. There might be pointed out two stages of the ERPT. Firstly, import prices respond to the change in the exchange rates, and only then do consumer prices respond.

Logically, ERPT affects import prices more comparing to consumer prices. This happens, as consumer prices take into account non tradable goods, which are not responsive to the exchange rate fluctuations. Interestingly, exchange rate shocks can affect prices at different stages (import prices, producer prices, consumer prices) directly and indirectly via the previous stages.

It should be also noted that pricing-to-market (PTM) is a closely related term, as it refers to the pricing behavior of exporters who export their goods to a destination market after the change in an exchange rate. Generally, pricing-to-market is defined as the percentage change in prices in the export's currency due to a one percent change in the exchange rate. For example, consider appreciation of the exporter's currency. In order to keep the price for a foreign market on the same level exporter is obliged to decrease that market's price expressed in the export's currency. If foreign market's prices (from the perspective of the exporter) in the export's currency will stay the same, this will lead to a complete pass-through effect. Thus, the lower is the effect of pricing-to-market the higher is the ERPT effect. It follows that if the exchange rate change is offset by the same proportional change in import prices, then the resulting pass-through effect is complete and hence there is no pricing to the market. At the other extreme, if exporter adjusts its prices expressed in its own currency diametrically proportional to the exchange rate change, then there will be a zero pass-through effect and full effect of pricing to the markets. This topic is further investigated in Goldberg and Knetter (1997), specifically the subject of study is the result of yen appreciation in the late 80-ies, which did not actually end up in the significant increase of the price (in dollars) of the Japanese products on the American market, as it might be suggested by the theory. Instead there was just a modest increase in the prices.

Pricing behavior of importing firms generally plays a huge role in the determination of the extent of the ERPT. There might be two options. Firstly, prices of the goods might be set in the importer's currency, then if no changes happen, the exchange rate fluctuations would be fully passed onto domestic import prices, thus there would be a complete ERPT.

Secondly, producers might price their good in the local currency (consumer's currency), then the exchange rate movements will not affect domestic prices, thus ERPT will be zero.

3.1 Incomplete ERPT Effect

In a literature on a given topic one can find several reasons to explain, why ERPT effect might be incomplete.

Firstly, as Krugman (1986) finds out, firms tend to price their goods according to the situation in the market, so they might want to adjust their profits just in order to keep their market share, instead of immediately changing the price. This may lead to international price discrimination and an incomplete pass-through in countries, where international arbitrage is difficult or maybe even impossible.

Rebelo (2002) argues that because of the fact that people change their preferences in favor of cheaper goods pass through effect is incomplete.

Also ERPT effect might be decreased by a common belief in reliable inflation stabilization by central banks. The low inflation is sure to change pricing behavior. Generally, the more stable a given county is and the lower its rate of inflation is, the lower the degree of ERPT will there be. As it often happens with inflation in the economic theory, this may lead to the virtuous cycle: country's high inflation leads to ERPT, then ERPT provokes even higher inflation and so on. This idea is supported by Campa and Goldberg (2004). This usually happens due to the fact that in the higher inflation environment, prices adjust more quickly.

Moreover, because of the increasing role of the globalization, international companies face much stronger competition. Firms no longer can pass costs connected with the change of the exchange rates onto prices. Thus, the degree of the pass-through effect also depends on the competition the exporter faces in a given market. For example, according to Feenstra (1993) certain industries face a high competition, which leads to a low ERPT and high PTM. This happens, for instance, on the automobile market and on the market of alcoholic beverages.

However, if the destination market's currency appreciates, it might be reasonable for exporter not to adjust its price, as goods then will become relatively cheaper. In such a case there will be a complete ERPT.

ERPT heavily depends on the type of goods mainly imported in the country, as different goods have different levels of the ERPT effect. For example, pass-through effect of the oil prices or energy is relatively high, whereas pass-through effect in the case of the manufactured products is quite low. Which means that country's aggregate pass-through depends on the import basket. Thus, "a move away from energy towards manufactured products will cause a decline in the aggregate pass-through to import prices" (Stulz, 2007). This in return will result in a lower ERPT to consumer prices.

Another variable, explaining the extent of the exchange rate pass through is the output gap, which is calculated as a difference between the real product and the potential product. Due to the fact, that at times, when output gap is positive, economy is running above its potential, demand is high and importers again in order to keep or gain market share need to keep competitive prices, mostly "can't afford" to react to the exchange rate changes, thus ERPT will not be full.

On the other hand, against incomplete ERPT speaks the fact that the constant changes of the price under the influence of the exchange rate fluctuations may negatively affect firm's image, its reputation. That is why exporters may not want to react to the fleeting exchange rate fluctuations. Exporters, however, are much more likely to react to the changes in the exchanges rates, that seem to be permanent.

3.2 Literature Overview

There is a vast amount of papers on a given subject, even besides the ones that were already mentioned. Vector autoregressive process is used in the majority of papers to analyze the impact of exchange rate movements on prices.

However, there are also some papers, in which single equations are used (see Campa and Goldberg (2005)).

But even if we concentrate precisely on the VAR approach to the pass-through modeling, the model is estimated in different ways. For example, in some of the papers dealing with VARs (Rowland, 2004) there are just 4 variables: exchange rate (EX), import price index (IMP), producer price index (PPI) and consumer price index (CPI). Instantaneous causal ordering of the variables then usually is $EX \rightarrow IMP \rightarrow PPI \rightarrow CPI$. Such papers generally find that pass-through effect is incomplete.

The most incomplete short-run ERPT is found in the paper by McCarthy (2007), ERPT is considered to be zero. McCarthy investigates the pass-through effect in the industrialized counties using variables such as oil prices, GDP gap, exchange rate, import prices, producer prices, consumer prices, broad money and an interest rate.

In a fundamental work by Mihailov (2005), which deals with the analyzing of the ERPT in set of countries, original Cholesky order of the variables is $money \rightarrow exchange\ rate \rightarrow import\ prices \rightarrow export\ prices \rightarrow inflation$. This order is suggested by Granger causality tests and is later modified in a following manner $inflation \rightarrow exchange\ rate \rightarrow money \rightarrow export\ prices \rightarrow import\ prices$.

Author suggests that this modification is possible due to several reasons:

- $money$ and $exchange\ rate$ can be switched, because "central bank policy pays some more attention (at least implicitly) to the exchange rate"
- Theory supports $inflation$ to be moved from last to first, as it is the "primary objective of most contemporary central banks".

Mihailov also finds out that the “orthogonal impulse response estimates of pass-through from the above four specifications have been relatively robust to ordering”.

There are also papers applying generalized impulse responses, where ordering of the variables does not matter. This approach can be found, for example, in Pesaran and Shin (1998).

Leigh and Rossi (2002) find out using VAR model that there is a large and short pass-through effect in Turkey, which mostly disappears after the first four months.

As papers suggest, ERPT is declining over time, ERPT is usually far from being complete in developed and developing countries (Likka Korhonen and Paul Wachtel (2005)). ERPT also was significantly higher in the 1980s than it is nowadays. Some of the reasons, such as the influence of the globalization, already mentioned in this paper, may explain this decline in the degree of ERPT. Less developed countries and countries with the vast presence of foreign firms, however, tend to have relatively high pass-through.

Returning to the Czech Republic, one might predict low ERPT onto CPI and a bit higher ERPT onto import prices. According to Babecká-Kucharčuková (2009) this is so due to several reasons. Firstly, “the Czech Republic is a small open economy with a ratio of exports/imports to GDP exceeding 60%” and also as “during 1999-2004 around 90% of contracts for imported goods were dominated in foreign currency”.

Importantly, ERPT might differ across countries, time periods, stages along the pricing chain, methods of estimation, time horizons and data frequencies. Also, as Stulz (2007) notices “evidence from VARs may heavily depend on the model specification”.

3.3 General Steps of Model Construction

In order to make it easier for the reader to understand the remaining parts of the paper, a brief description of the models creation will be provided here.

1. We will search the data for the variables, which seem to have an economic rationale for the inclusion into the model.
2. We will apply unit root tests on the seasonally adjusted variables.
3. Granger causality between the variables then will be examined.
4. We will search for the cointegrated relationships between the variables.
5. To find an optimal VAR order we will address to the information criteria.
6. Different dummy variables will be used, so that the model passes the normality tests and the autocorrelation test. After that, we will again look at the information criteria.
7. We will examine original, as well as Cholesky accumulated impulse responses.
8. And, finally, we will interpret the findings of the 7th step.

3.4 Variables in the Model

The main purpose of this paper is to measure the effect of the volatility of the exchange rate onto consumer prices, so these are the first two variables to be included into the model.

Monthly observations will be taken into account, because according to Mihailov (2005) “pass-through has to do with reactions of monopolistically competitive price-setters to (i) exchange rate movements (ii) under sticky prices. On both counts, quarterly observations would miss much of the “action””.

Nominal effective exchange rate (EX) is in the form of index and is calculated as geometrical weighted average of the nominal exchange rates of individual countries, where weights correspond to the import and export shares of those counties, which are considered to be the largest trading partners of Czech Republic. Appreciation of the Czech koruna is represented by the higher values of index; depreciation in return is represented by lower values of the index. Alternatively, inversed nominal effective exchange (iEX) rate could have been used in the model. It is calculated as $iEX = {}^{1}/_{EX}$.

Source: IMF, monthly data. 2005=100.

Import prices index (IPI). Prices are without taxes. The increase in the import prices index represents the percentage increase of a mean import price level comparing it with the mean import price level of the corresponding period of the previous year.

Source: Czech Statistical Office, monthly data. 2005=100.

Consumer price index (CPI) characterizes the average price development in the country. Average 2005=100.

Source: IMF, monthly data.

USA consumer price index (USA) analogously represents the price development in America.

Source: IMF, monthly data.

Economic activity is also included into the model. It is measured as the *output GAP*, which is calculated as a difference between the real product and the potential product. Of course, firstly we need to get GDP data. It is quarterly available on the web pages of the Czech Statistical Office (CZSO). But due to the fact that monthly datasets are to be used in the model, it is needed to recalculate quarterly GDP data using EViews software. Another option is to apply Chow-Lin procedure, which is an effective procedure of transforming low frequency series into higher frequency series. However, as EViews definitely provides an easier way of getting the similar results, it is going to be used (see Frequency conversion: quarterly to monthly). Later, those monthly observations would be recalculated using Hodrick-Prescott filter. It is a commonly used approach of creating GAP series.

Oil prices (oil) represent the movement of oil prices. Calculated in the US dollars per barrel.

Source: IMF, monthly.

Producer prices (PPI) measures the average change over time in the selling prices received by domestic producers for their output.

Source: IMF, monthly.

3.5 Frequency Conversion: Quarterly to Monthly

EViews offers several methods for the data transformation. But not all of the offered methods can provide equally satisfactory results. However, quadratic-match average seems to be the best option for the papers purposes. Basically, quadratic-match average performs a local quadratic interpolation of the three adjacent points from the source series, so that their average matches the low frequency data actually observed.

For example, consider following GDP low frequency data actually used in the model:

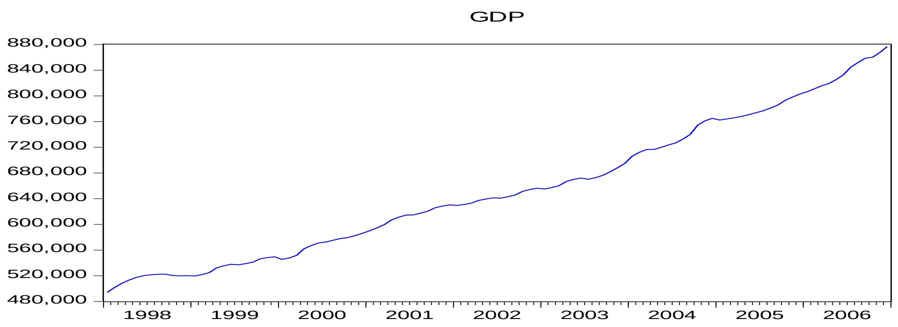
2002Q1	631200
2002Q2	639426

They are to be transformed into:

2002M01	627647.2592
2002M02	631362.1481
2002M03	634590.5925
2002M04	637332.5925
2002M05	639588.1481
2002M06	641357.2592
2002M07	640972.2222
2002M08	643019.2222

The first three monthly observations have an average of 631200. To create higher frequency data one should estimate parameters A, B and C of $At^2 + Bt + C$, where t is a trend, keeping in mind that the average for the first three observations should be 631200, for the later three the average should be equal to 639426 etc.

Below please find a graph of monthly observations of GDP:



Graph 3-1: Monthly GDP observations

3.6 Estimation, Tests, Results

The sample is chosen in such a way that the years of neither Asian economic, nor of the global crisis would not be included into the model. We also want to keep an option of making a prediction for the non-crisis years of 2005 and 2006. Larger sample size was also a subject of consideration, but due to the structural changes in the data, it was needed to use a lot of dummy variables and moreover the final model could have been incorrect. So the final decision is in a favor of smaller, but “stable” time interval of the years from 1999 till 2004. Later to check the adequacy of the model a forecast for following two years will be made and then compared with the actual data.

Now we are going to study the relationships between the inflation in the USA, oil prices (both of them are exogenous) and the endogenous domestic inflation and exchange rate. In other words, we assume that the domestic inflation and the exchange rate affect one another.

All of the variables will be seasonally adjusted by the Census X12 procedure. It is also necessary to test every variable for a unit root. One should firstly address to the ADF test statistic, then to the KPSS statistic and finally Phillips-Perron test can be taken into account. All of the variables are shown to be $I(1)$ by all of the tests. Details can be found in the appendix. One of the possible ways of dealing with the non-stationarity $I(1)$ is to take the first differences.

One might also want to test for the Granger causality between the variables, but it is important to keep in mind that there are some problems associated with this method. Firstly, in a way it simplifies the problem. Moreover, one can never be sure that the considered time horizon is large enough; on the other hand too large sample period may hide causality. Also according to Sorensen (2005) “It will very often be hard to find any clear conclusions unless the data can be described by a simple 2-dimensional" system”. After all, Granger causality test does not provide any concrete results, but it might suggest that the use of some variable is in a way justified.

Let’s now address to the data under consideration. To start with, stationary variables should be used in the Granger causality test. The Granger causality findings correspond with the earlier mentioned cons of the given test. Exchange rate is found to Granger cause domestic inflation in the sample from 1996 till 2006 (once more, such a big sample was not used in the model due to economic crises and because we wanted to make a prediction), but no Granger causality was found in a smaller period which is later used in the model. Although, no Granger causality between other variables was found, they seem to have clear economic rationale to be included into the model.

Inflation in America represents global factors affecting the level of inflation in the Czech Republic. So, by including, this variable we hope not to allow the mistake of omitting an important variable in the model specification to happen. We assume that the American CPI index will provide our model with an “average influence of the omitted variables”. We could have alternatively used inflation in the euro area. Oil prices to some extent may then be explaining price and exchange rate fluctuations. Generally, oil prices identify supply shocks.

If series are integrated of the same order, we need to test for the possible existence of cointagrated relationships. In the case of cointegration, achieving stationarity of series by using first differences will not be correct. More precisely, then vector error correction (VEC) model should be used. Before coming to the explanations of the test, we might assume that the given variables are somewhat "between having and not having a cointagrated relationship", as some of the researchers instead of VARs prefer using VECs. We are going to examine cointegration using Johansen test. Non-stationary series should be used. No cointegration is found between different pairs of the variables. For the detailed test output please address to the appendix. We found that the use of a VAR model is justified in a given case.

Our four-variable VAR model will consist of the first difference of the oil prices, the first difference of the domestic consumer prices, the first difference of USA prices and the first difference of the koruna's nominal exchange rate.

So, in a result, $y_t = [\Delta EX, \Delta CPI]$. Since the Czech Republic is a small economy, the first difference of the oil prices and the first difference of USA prices would be exogenous variables.

It seems reasonable to let oil prices and USA CPI have an immediate as well as lagged effect on both of the endogenous variables. In a current study non-lagged and lagged by one period oil prices and USA CPI index are used.

Vector autoregression of order two (VAR(2)) is a priori estimated. Then we address to the informational criteria. By a coincidence final prediction error and Akaike informational criterion suggest using exactly two lags. This also seems to be reasonable due to our relatively small sample. Output is again in the appendix.

Later one should look at the residuals in order to handle the outliers with dummy variables. Usually such outliers have a clear economic explanation. We use just one dummy variable for the period of 1999M02. This might be connected with the introduction of euro in the January 1999. Afterwards, dummy is included into the VAR model as exogenous variable. One more check of the residuals is than needed. This time no extreme outliers are found. Inverse roots of the characteristic polynomial are found to lie within the unit circle, thus the given VAR process is stable.

We might continue with the studying of the normality of the model. Residuals are found to be multivariate normal by the normality test. This result is crucial for the further use of the model.

No less important is the possible problem of the autocorrelation. It is investigated with the help of correlograms and autocorrelation tests. No statistically significant autocorrelation is found.

Final CPI equation has a form of:

$$cpi_t = 0.023 + 0.0008ex_{t-1} - 0.034ex_{t-2} + 0.21cpi_{t-1} + 0.29cpi_{t-2} + \\ +0.012oil + 0.024oil_{t-1} + 0.127usa + 0.095usa_{t-1} + 0.087dummy$$

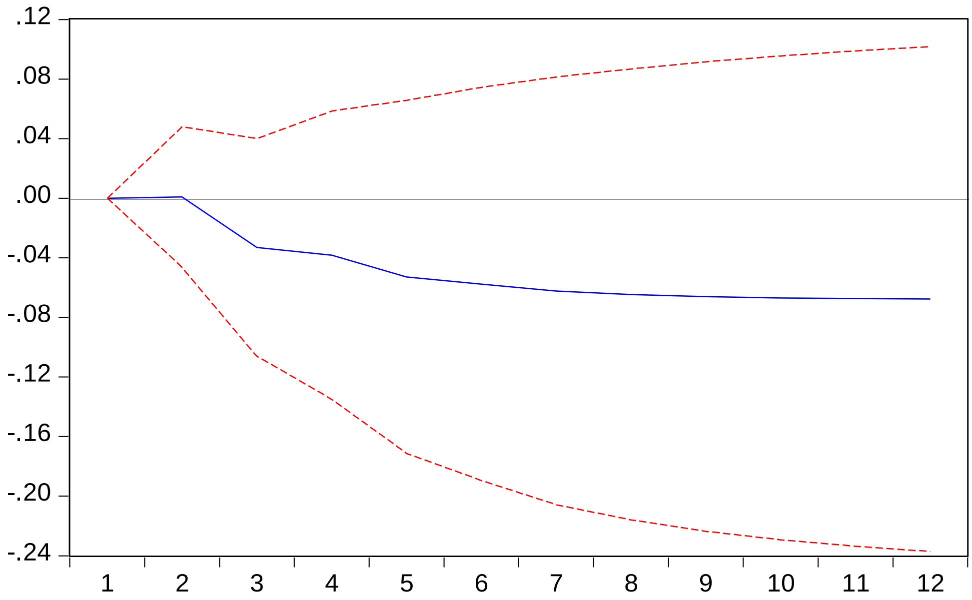
Where lower case letters indicate the first differences of the variables.

Due to the fact, that the most important tests are passed, we might continue with the actual analysis of the pass-through effect. In our particular study it seems to be the most reasonable to analyze accumulated impulse responses. Accumulated impulse responses at time horizon T is just a sum of all the impulses from the time 0 till T.

As shocks in the exchange rate and in the domestic inflation have a small correlation (see residual correlation matrix, as always, in the appendix), there will not be a big difference in using orthogonal Cholesky impulses or just ordinary ones with the unchanged covariance matrix.

According to Eduardo Rossi (lecture notes, published on the internet) "if the variables have different scales, it is sometimes useful to consider innovations of one standard deviation rather than unit socks".

Let's now discuss and compare both of the options: accumulated response of the domestic inflation to a nonfactorized one unit shock in the exchange rate and accumulated response of the domestic inflation to a Cholesky one standard deviation shock in the exchange rate series.



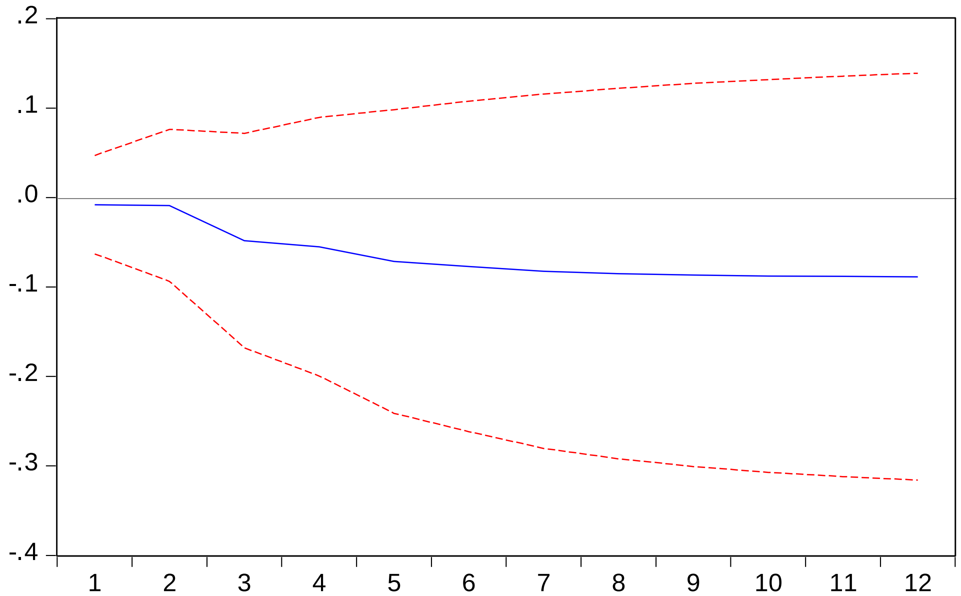
Graph 3-2: Accumulated responses of the domestic inflation to a one unit shock in the exchange rate

Where the blue line represents accumulated impulse responses and the red dotted line stands for the response standard errors. Firstly, we note that, as nonfactorized impulse responses assume that shock in one variable is not accompanied by a shock in another variable, exchange rate shock is passed to the domestic inflation only in the second period. It will not be the case, when Cholesky decomposition (shocks) will be used.

In the second period we observe insignificant shock of 0.0008, equal to the coefficient of the ex_{t-1}. Accumulated shocks in the third period are -0.033, such a step decline happens due to a much lower coefficient of ex_{t-2} (-0.034) than $ex_{t-1}(0.0008)$. Similarly, we continue till the 12th month, when the accumulated shocks "stabilize" at the level of -0.068.

Previous method does not consider for the possible correlation between the shocks of different variables. Now let's continue with the orthogonal Cholesky impulses (basically, recursive VAR impulse responses). As already have been mentioned, it is expected to get quite similar results, as correlation between the exchange rate and the domestic inflation is low.

Due to instantaneous causal ordering of the variables $(ex \rightarrow cpi)$ there will be instantaneous effect of the exchange rate on the domestic inflation (see chapter on the Impulse Response Analysis in the theoretical pat of the paper). Also, according to Rossi (lecture notes) and to the theoretical chapter, "the variances of the components are one. Thus, a unit innovation is just an innovation of size one standard deviation."



Graph 3-3: Accumulated responses of the domestic inflation to a Cholesky one standard deviation innovation in the exchange rate

Immediate effect of the one standard deviation (alternatively, one unit) shock of the exchange rate on the Czech CPI is -0.01. In the third month the accumulated impulse responses are -0.05. In the sixth period they equal to -0.08 and, finally, in the eighth month accumulated impulse responses "stabilize" at the level of -0.09.

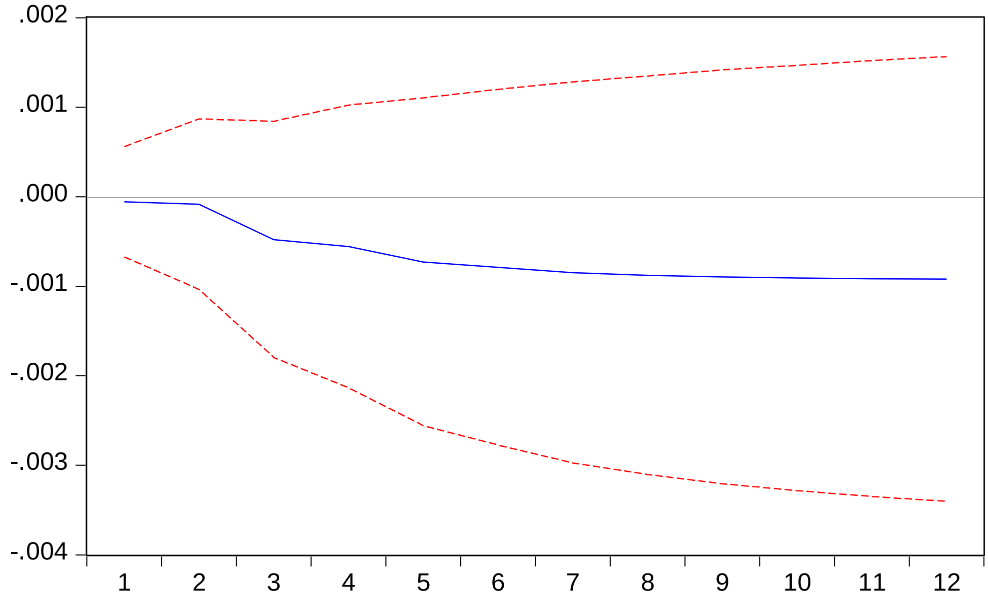
We get different accumulated impulse responses due to the distinctions in the manipulations with the covariance matrix. Accumulated impulse responses based on the

Cholesky decomposition are lower due to the negative correlation between the residuals of the nominal Czech koruna's exchange rate and the domestic CPI.

As impulse responses based on the Cholesky decomposition better describe the reality, we conclude that appreciation of the Czech koruna equal to 1% increase in the corresponding nominal effective exchange rate results in the 0.01% domestic deflation after one month, 0.05% deflation after three months and, finally, in a long run it stabilizes at the mark of 0.09% deflation. These coefficients also may be considered as the degrees of the ERPT effect in different months. Under the same logic, depreciation of the Czech koruna equal to 1% decrease in the corresponding nominal effective exchange rate results in the 0.01% domestic inflation after one month, 0.05% inflation after three months and, finally, in a long run stabilizes at 0.09% inflation.

These findings do correspond with the theoretical assumptions and with other papers on the given subject.

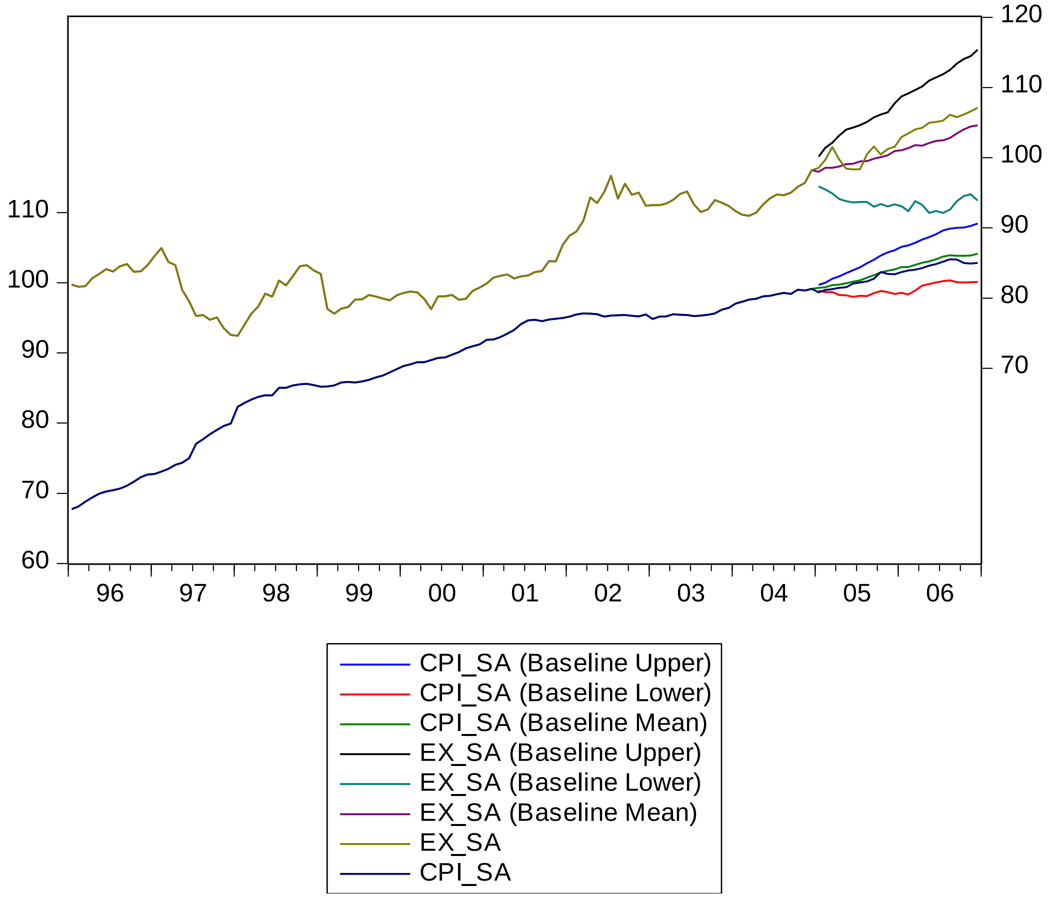
As possible critiques of the model two points may come into one's mind. Firstly, the variables of the model are not in the form of natural logarithms. This decision, however, was based on the fact that the min and max values of time series do not differ much. Moreover, time series not necessary show the exponential growth. After all, results will be the same, in a given case, it is just a matter of rescaling. Below please find accumulated impulse responses for the model with the same variables taken as natural logarithms:



Graph 3-4: Accumulated responses of the domestic inflation to a Cholesky one standard deviation innovation in the exchange rate in a logarithmic scale

The second point may be that in the literature on a given subject much more frequent approach is to take inversed nominal effective exchange rate. However, it will just slightly change the interpretation of impulse responses (as they will be inversed).

To check the adequacy of the model a forecast for a following two non-crisis years will be made. As we can see from the graph below, predictions of the model for the period from 2005m01 till 2006m12 match actual movements of the variables under consideration very well. Thus, we conclude that the model somewhat "catches" the actual relationships between the variables and is close to the reality.



Graph 3-4: Forecast for a following two years

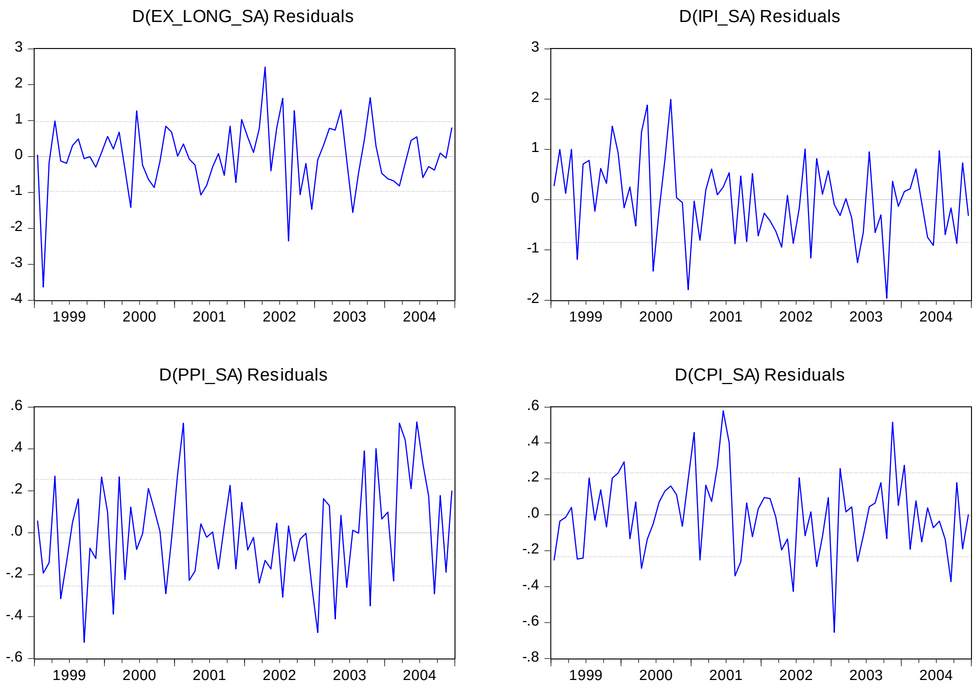
3.7 Alternative Specification

In order to check different possible specifications of the model and to study the distribution chain, where exchange rate shocks are mainly firstly passed to the import prices, then to the producer prices and, finally, to the consumer prices, another model in a spirit of McCarthy (2007) will be discussed in the following section. Please note, that some technical issues will not be described in the details, as it was in the case of the first model.

There will be four endogenous variables: nominal effective exchange rate, import prices, producer prices and consumer prices. Exogenous supply shock will be represented by oil prices, demand shock by output gap. Both of them are sure to influence the response of prices to the exchange rate fluctuations. Exogenous variables are lagged by one period. The same data sample of the years 1999M01 till 2004M12 will be used.

There is no need to compute unit root tests for the nominal effective exchange rate, domestic CPI and oil prices, as it has already been done in the previous model. Those variables are shown to be $I(1)$. Output gap is stationary by its construction. Tests for the remaining variables please, as usual, find in the appendix of the paper. We should proceed by taking the first differences of all the variables, except the output gap. Later, we test for a cointegration among the variables of the model. No cointegrated relationships among different pairs of the variables were found.

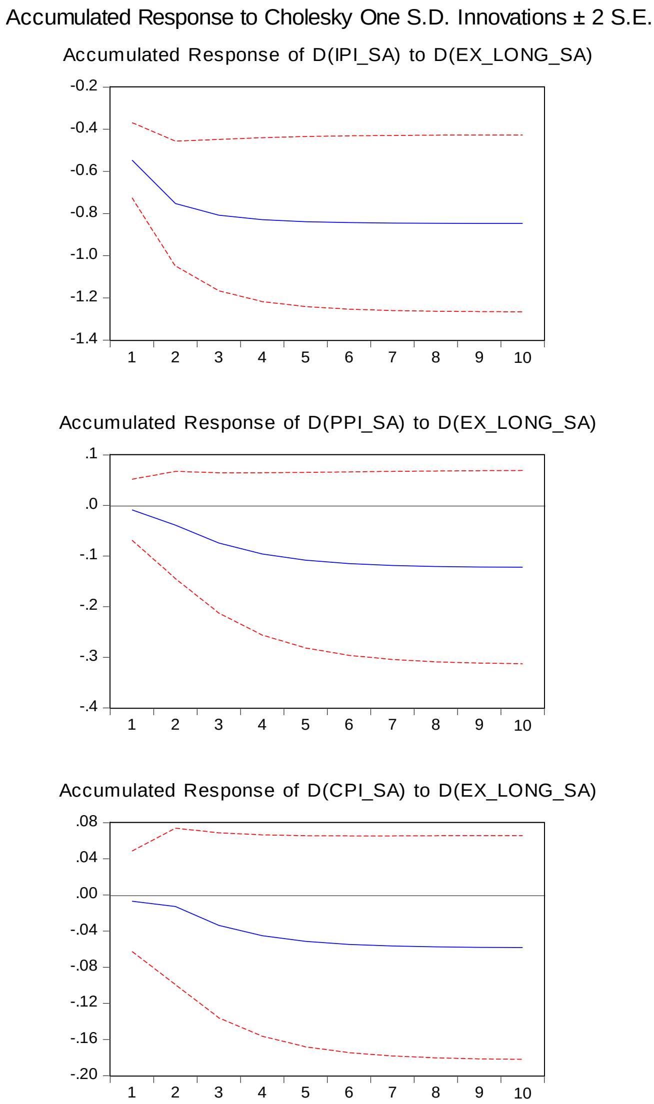
Then we a priori estimate a VAR of order two, after that, following a suggestion of the information criteria, we estimate VAR(1). We look at the normality tests. By far, it is not passed. Thus, we address to the graphs of the residuals:



Graph 3-5: Residuals of the alternative specification

We note that there is an outlier at the February of the year 1999. We handle it with a dummy variable. Then look at the normality test. This time it is passed. Portmanteau test for the autocorrelations does not find any significant autocorrelations among the residuals. Inverse roots of the characteristic polynomial are found to lie within the unit circle, thus given VAR process is stable.

We continue with the impulse response functions. We study the effect of the Cholesky orthogonal shock in the exchange rate on the prices along the distribution chain (import, producer and consumer prices).



Graph 3-6: Accumulated impulse responses to a Cholesky one standard deviation innovation in the exchange rate

The smallest and the slowest ERPT, as predicted, is to the consumer prices. After ten months it is 0.058. Producer prices react more to the change in the exchange rate. Their ERPT is estimated to be 0.12 after ten months. Import prices react immediately and after ten months they show a vast ERPT effect of 0.85.

These findings are in line with the assumption that ERPT declines along the distribution chain. The adjustment speed also decreases.

As a possible critique of both of the models one might mention the fact that we did not consider for the effects of the monetary policy. They could have been represented by the short-run interest rate. Such variable was included into the last model, the results were more or less similar. Generally, ERPT showed to be a bit higher for all of the prices.

3.8 Brief Recapitulation of Methods of Model Construction

We have chosen the variables of the model in the consistency with the economic theory and logic. We have gained all the relevant data from different internet sources. We have constructed monthly GDP series in order to transform it to GAP series. Then we found that the majority of the seasonally adjusted variables are integrated of the first order. It caused the need to test for the cointegrated relationships. However, no cointegration was found. It "justified" the use of the VAR framework. We took the first differences for all the variables, except GAP, which is stationary by its construction. We a priori estimated VAR of the second order. After that, we estimated VAR of the consistent with information criteria order. We have achieved normality and zero autocorrelations among the residual series by using dummy variables. Results of the information criteria are much more trustworthy under the assumptions of the normality and zero autocorrelation mentioned earlier, so we have checked them once more. For the actual analysis of the pass-through effect we have used different types of impulse responses. However, Cholesky orthogonal impulses better match the purposes of the given paper. Both of the considered models showed similar results, that the exchange rate pass-through to the consumer prices was at the level of $-0.06 - 0.09$ in the beginning of the 21st century in the Czech Republic.

4 Conclusion

Theoretical part of the paper aims to provide a clear and at the same time accurate discussion on the most important topics connected with the vector autoregression. We discuss, for example, different representations of the vector autoregressive processes, stability conditions, Granger causality and impulse response functions, which are later used in the second part devoted to the analysis of the exchange rate pass-through.

An extremely complex pass-through theory may provide with the various possible specifications of the model. However, only two of them were chosen for the impulse response analysis in the given paper.

First specification deals with just four variables, two of them are endogenous: nominal effective exchange rate of the Czech koruna and the domestic CPI index. Two exogenous variables (oil prices and the CPI index in the USA) represent supply shocks and global factors that can to some extent affect the level of the domestic inflation. Both exogenous variables are allowed to have lagged values. Generally, use of such variables is not in the original spirit of the vector autoregressive models, but it would be economic non-sense to allow the inflation level or the exchange rate of the relatively small economy to have an influence on the worldwide inflation or oil prices. As already been mentioned, the main source of the analysis is impulse response functions. We also use accumulated impulse responses; it makes possible to "catch" the extent of the exchange rate pass-through effect and its speed. Both of the general one unit shocks and Cholesky orthogonal shocks are used. Due to a low correlation among the residual series of the endogenous variables, results are quite similar, otherwise they would not be. We find that domestic prices react to a 1% change in the nominal effective exchange rate by 0.09% after eight months. We also observe quite a fast reaction. Finally, in order to check the adequacy of the model we make a prediction for a following two years period and then compare it with the actual data movements. Prediction happens to be nearly perfect. Thus, we conclude that the first model is representing real relationships between the variables.

Second specification follows some steps of McCarthy's works. It examines pass-through effect via so called distribution chain, where exchange rate shocks are mainly firstly passed to the import prices, then to the producer prices and, finally, to the consumer prices. However, direct impact is also allowed. Current specification, thus, have four endogenous variables: nominal effective exchange rate, import prices, producer prices and consumer prices. We will also consider a bit different composition of the exogenous variables. Exogenous supply shock will be represented by oil prices, demand shock by output gap. Output gap is included under the logic that, at times when it is positive, economy is performing incredibly well, demand is high and importers mostly cannot react to the exchange rate fluctuations, because otherwise they would not have competitive prices; thus, exchange rate pass-through will not be full. Both of the exogenous variables are sure to influence the response of prices to the exchange rate fluctuations. They are also allowed to have lags. We find that import prices react immediately to the changes in the exchange rate. ERPT in such a case is considered to be 0.85% after ten months for every percentage change in the nominal effective exchange rate. Producer prices then react much slower. ERPT effect for the producer prices after ten months is just 0.12%. We observe the lowest and the slowest ERPT in the case of the domestic consumer prices. It is counted to be 0.058% after ten months. Theoretical assumption of ERPT affecting the import prices more comparing to the consumer prices is thus fulfilled.

5 Literature

Books used for the part of the paper devoted to the vector autoregression theory listed in the alphabetical order:

[1] Cipra, T.: *Finanční Ekonometrie*. Ekopress, 2008. ISBN 978-80-86929-43-9.

[2] Enders, W.: *Applied Econometric Times Series*. Wiley Series in Probability and Statistics, 2009. ISBN 978-0470505397.

[3] Gujarati, D.: *Basic Econometrics*. McGraw Hill, 2008. ISBN 978-0073375779.

[4] Hušek, R.: *Ekonometrická Analýza*. Oeconomica, 2007. ISBN 978-80-245-1300-3.

[5] Lütkepohl, H.: *New Introduction to Multiple Time Series Analysis*. Springer, 2005. ISBN 3-540-40172-5.

[6] Rachev, S.: *Financial Econometrics from Basics to Advanced Modeling Techniques*. John Wiley & Sons, 2007. ISBN 978-0-471-78450-0.

[7] Tsay, R.: *Analysis of Financial Time Series*. John Wiley & Sons, 2005. ISBN 13 978-0-471-69074-0.

Sources used for the practical part of the paper; again in the alphabetical order:

[1] Babecká-Kucharčuková, O.: Transmission of Exchange Rate Shocks into Domestic Inflation: the Case of the Czech Republic. *Czech Journal of Economics and Finance*, 2009, vol. 59, No. 2, pages 137-152.

[2] Campa, J.: Exchange Rate Pass-Through into Import Prices [online]. *Federal Reserve Bank of New York*, 2004. At http://www.newyorkfed.org/research/economists/goldberg/passthrough093004.pdf.

[3] Feenstra, C.: Market Share and Exchange Rate Pass-Through in World Automobile Trade [online]. *Federal Reserve System International Finance Discussion Papers*, 1993, No. 446. At http://www.c.federalreserve.gov/pubs/ifdp/1993/446/ifdp446.pdf.

[4] Goldberg, P.: Good Prices and Exchange Rates: What Have We learned? *Journal of Economic Literature*, 1997, vol. 35, No. 3, pages 1243-1272.

[5] Korhonen, L.: A Note on Exchange Rate Pass-Through in CIS countries [online]. *Bank of Finland Discussion Papers*, 2005, Vol. 2. At http://www.suomenpankki.fi/pdf/118266.pdf.

[6] Leigh D.: Exchange rate pass-through in Turkey [online]. *IMF Working Paper*, 2002, 02/204. At http://www.imf.org/external/pubs/ft/wp/2002/wp02204.pdf.

[7] McCarthy, J.: Pass-Through of Exchange Rates and Import Prices to Domestic Inflation in Some Industrialized Economies. *Eastern Economic Journal*, 2007, Vol. 33, No. 4.

[8] Mihailov, A.: Exchange Rate Pass-Through on Prices in US, German and Japanese Macrodata [online], 2005.
At http://www.univ-orleans.fr/deg/GDRecomofi/Activ/mihailov_strasbg05.pdf.

[9] Pesaran, M.: Generalised Impulse Response Analysis in Linear Multivariate Models. *Economics Letters*, 1998, vol. 58, pages 17-29.

[10] Rebelo, S.: Why are Rates of Inflation So Low after Large Devaluations? [online] *Rochester Center for Economic research Working Paper No. 486*, 2002.
At http://rcer.econ.rochester.edu/RCERPAPERS/rcer_486.pdf.

[11] Rossi, E.: Lecture Notes on the Impulse Response Functions [online]. Universita di Pavia.
At http://economia.unipv.it/pagp/pagine_personali/erossi/dottorato_svar.pdf.

[12] Rowland, P.: Exchange Rate Pass-Through to Domestic Prices: the Case of Colombia [online]. Banco de la Republica Colombia, 2004, No. 47.
At http://www.banrep.org/docum/ensayos/pdf/espe_047-3.pdf.

[13] Sorensen, B.: Lecture Notes on the Granger Causality [online]. 2005.
At http://www.uh.edu/~bsorense/gra_caus.pdf.

[14] Stulz. J.: Exchange rate pass-through in Switzerland: Evidence from vector autoregressions [online]. *Swiss National Bank Economic Studies*, 2007, No. 4.
At http://www.snb.ch/n/mmr/reference/economic_studies_2007_04/source.

[15] Wikipedia http://en.wikipedia.org/wiki/Bayesian_information_criterion, [cit. 29.12.2012].

6 Appendix

6.1 Unit Root Tests

Null Hypothesis: EX_SA has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.656980	0.2574
Test critical values:	1% level	-4.092547	
	5% level	-3.474363	
	10% level	-3.164499	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(EX_SA)
Method: Least Squares
Date: 12/23/12 Time: 19:59
Sample (adjusted): 1999M02 2004M12
Included observations: 71 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
EX_SA(-1)	-0.135830	0.051122	-2.656980	0.0098
C	10.38947	3.949118	2.630833	0.0105
@TREND(1999M01)	0.046977	0.016184	2.902708	0.0050

R-squared	0.110274	Mean dependent var	0.208041
Adjusted R-squared	0.084105	S.D. dependent var	1.220502
S.E. of regression	1.168050	Akaike info criterion	3.189883
Sum squared resid	92.77511	Schwarz criterion	3.285489
Log likelihood	-110.2408	Hannan-Quinn criter.	3.227902
F-statistic	4.213997	Durbin-Watson stat	1.774944
Prob(F-statistic)	0.018824		

Evidences from the upper table suggest, that the t-Statistic (-2.6569) in the case of the nominal exchange rate is higher than the corresponding critical value at 5% significance level (-3.474363), thus we cannot reject H_0 that there is a unit root, therefore monthly EX_SA observations are $I(1)$. One should also always follow the Durbin-Watson statistic, otherwise current test may not be reliable. In a given example its value (1.774944) belongs to a $< 1.72, 2{,}28 >$ interval, it suggests that there is no autocorrelation.

We continue with the Phillips-Perron unit root test:

Null Hypothesis: EX_SA has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 4 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-2.846756	0.1861
Test critical values:	1% level	-4.092547	
	5% level	-3.474363	
	10% level	-3.164499	

As the p-value is 0.1861, we can't reject the hull hypothesis of EX_SA having a unit root neither on the 0.05 probability, nor 0.1.

Then we can address to the Kwiatkowski-Phillips-Schimdt-Shin test. The corresponding null hypothesis this time is different. The null hypothesis for the KPSS test is stationarity.

Null Hypothesis: EX_SA is stationary
Exogenous: Constant, Linear Trend
Bandwidth: 6 (Newey-West automatic) using Bartlett kernel

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.123963
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
	10% level	0.119000

At the five percent significance level nominal exchange rate of Czech koruna is found to be stationary, but due to the fact that it is maybe in a given case better to use an option of only an intercept in the test equation (instead of trend and intercept) and because other tests find EX_SA to be $I(1)$, we may conclude that the nominal exchange rate of Czech koruna is found to be non-stationary.

Null Hypothesis: EX_SA is stationary
Exogenous: Constant
Bandwidth: 6 (Newey-West automatic) using Bartlett kernel

		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		1.012906
Asymptotic critical values*:	1% level	0.739000
	5% level	0.463000
	10% level	0.347000

We should repeat the following steps for all of the variables and, in the case that they are not stationary, for their first differences too. To avoid repetitions, all the results of the tests are summarized in the table below:

Variable	ADF		PP		KPSS	
	Level	1.Diff	Level	1.Diff	Level	1.Diff
ex_sa	-2.656980	-9.781674	-2.846756	-9.675004	0.123963	0.186544
cpi_sa	-1.654149	-3.642500	-1.235333	-6.879921	0.231513	0.263694
oil_sa	-2.505693	-9.002147	-2.450317	-9.037940	0.715064	0.093083
cpi_usa_sa	-1.699842	-7.944661	-1.821168	-7.149404	0.149350	0.109417

This table suggests that all of the variables are difference stationary.

6.2 Granger Causality Tests

Pairwise Granger Causality Tests
Date: 12/23/12 Time: 21:51
Sample: 1996M01 2006M12
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
D(EX_SA) does not Granger Cause D(CPI_SA)	129	5.58774	0.0047
D(CPI_SA) does not Granger Cause D(EX_SA)		0.84206	0.4333

Nominal Exchange rate of the Czech koruna Granger causes domestic inflation on the 5% significance level in the sample period from 1996 till 2006.

Pairwise Granger Causality Tests
Date: 12/23/12 Time: 22:07
Sample: 1999M01 2004M12
Lags: 4

Null Hypothesis:	Obs	F-Statistic	Prob.
D(CPI_SA) does not Granger Cause D(EX_SA)	67	0.73692	0.5706
D(EX_SA) does not Granger Cause D(CPI_SA)		0.81642	0.5199

In a another sample period (1999M01 2004M12), nominal Exchange rate of the Czech koruna does not Granger cause domestic inflation There is no point in showing all the combinations of the variables with different lags. For such a reason, just the last test will be demonstrated:

Pairwise Granger Causality Tests
Date: 12/23/12 Time: 23:18
Sample: 1999M01 2004M12
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
D(CPI_SA) does not Granger Cause D(CPI_US_SA)	69	0.07709	0.9259
D(CPI_US_SA) does not Granger Cause D(CPI_SA)		0.13720	0.8721

6.3 Johansen Cointegration Tests

No cointagration is found between different pairs of the variables. Several parings will be showed below.

Series: EX_SA CPI_SA
Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None	0.137620	12.46422	15.49471	0.1360
At most 1	0.032057	2.248156	3.841466	0.1338

Trace test indicates no cointegration at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None	0.137620	10.21606	14.26460	0.1981
At most 1	0.032057	2.248156	3.841466	0.1338

Max-eigenvalue test indicates no cointegration at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

There is no stationary combination of the nominal exchange rate series and the CPI index series.

The same is true for the USA CPI index and the nominal exchange rate of the Czech koruna:

Series: CPI_US_SA EX_SA
Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None	0.046579	3.332797	15.49471	0.9497

At most 1	0.000603	0.041586	3.841466	0.8384

Trace test indicates no cointegration at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Domestic CPI and USA CPI also show no cointegration:

Series: CPI_US_SA CPI_SA
Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None	0.110120	8.233602	15.49471	0.4408
At most 1	0.002655	0.183433	3.841466	0.6684

Trace test indicates no cointegration at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

6.4 Selection of VAR Lag Order

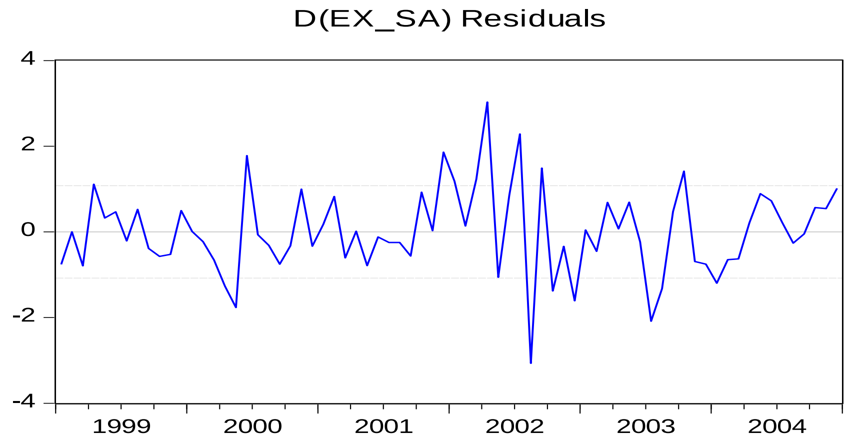
VAR Lag Order Selection Criteria
Endogenous variables: D(EX_SA) D(CPI_SA)
Exogenous variables: C D(OIL) D(CPI_US_SA(-1)) D(CPI_US_SA) D(OIL(-1))
Date: 12/24/12 Time: 20:09
Sample: 1999M01 2004M12
Included observations: 72

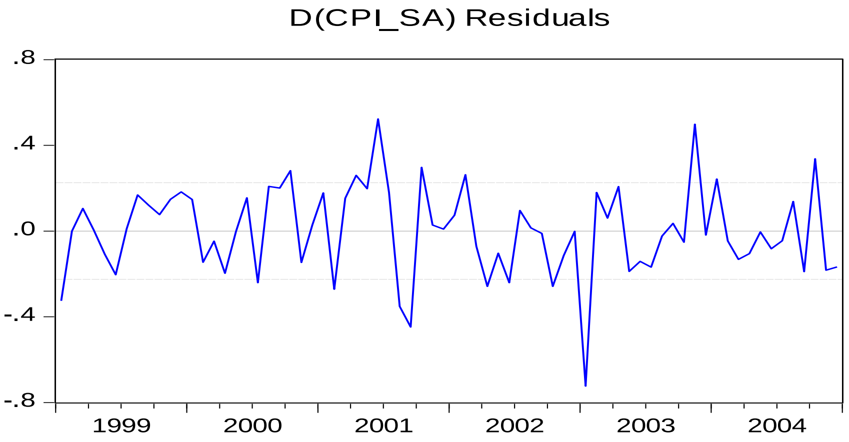
Lag	LogL	LR	FPE	AIC	SC	HQ
0	-112.9086	NA	0.104228	3.414128	3.730332*	3.540010*
1	-109.4439	6.255634	0.105872	3.428999	3.871684	3.605233
2	-101.7262	13.50599	0.095618*	3.325729*	3.894896	3.552316
3	-100.1634	2.648195	0.102540	3.393427	4.089075	3.670367
4	-99.17141	1.625713	0.111831	3.476984	4.299113	3.804276
5	-96.50825	4.216672	0.116571	3.514118	4.462729	3.891763
6	-89.68350	10.42670*	0.108399	3.435653	4.510745	3.863650
7	-84.81935	7.161112	0.106616	3.411649	4.613222	3.889999

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

VAR of order two is selected. This option is consistent with the FPE, the AIC and the fact, that the sample period is relatively small.

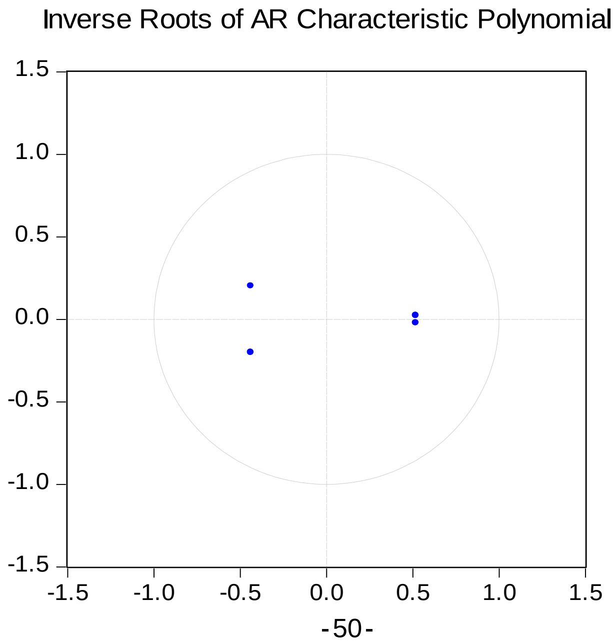
6.5 Graph of the Residuals





It is possible to conclude that after using one dummy variable, there are no extreme residuals left.

6.6 Inverse Roots of the Characteristic Polynomial



Inverse roots of the characteristic polynomial are found to lie within the unit circle, thus given VAR process is considered to be stable.

6.7 Normality Test

VAR Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: residuals are multivariate normal
Date: 12/24/12 Time: 22:37
Sample: 1999M01 2004M12
Included observations: 72

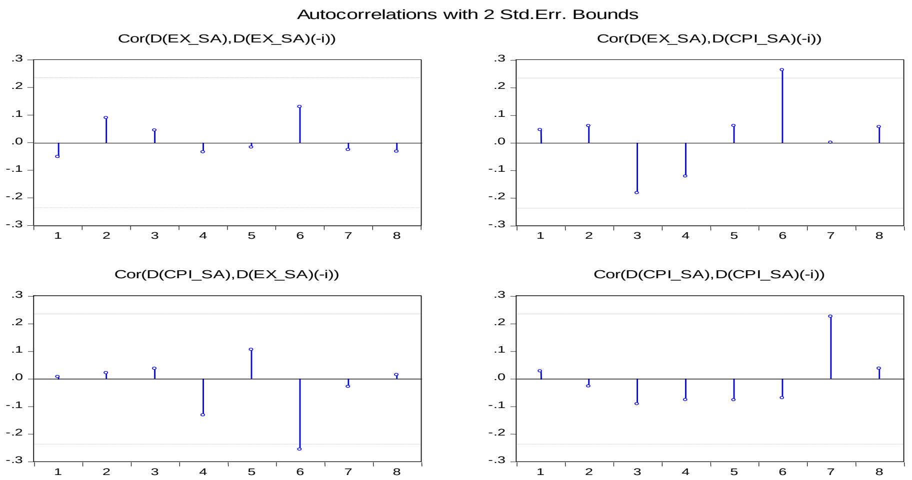
Component	Skewness	Chi-sq	df	Prob.
1	0.102309	0.125606	1	0.7230
2	-0.314555	1.187338	1	0.2759
Joint		1.312944	2	0.5187

Component	Kurtosis	Chi-sq	df	Prob.
1	4.051864	3.319252	1	0.0685
2	4.081643	3.509854	1	0.0610
Joint		6.829107	2	0.0329

Component	Jarque-Bera	df	Prob.
1	3.444858	2	0.1786
2	4.697192	2	0.0955
Joint	8.142050	4	0.0865

As this test suggests residuals of the investigated vector autoregressive model are found to be multivariate normal.

6.8 Correlograms and Autocorrelation Tests



Although, correlograms do testify in favor of possible autocorrelation of the order six, corresponding bounds are just slightly crossed and Portmanteu test suggests that there is no residual autocorrelation in the specified model. So it is possible to conclude, that no statistically significant autocorrelation was found.

VAR Residual Portmanteau Tests for Autocorrelations
Null Hypothesis: no residual autocorrelations up to lag h
Date: 12/24/12 Time: 23:37
Sample: 1999M01 2004M12
Included observations: 72

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	0.394924	NA*	0.400486	NA*	NA*
2	1.351782	NA*	1.384683	NA*	NA*
3	4.632993	0.3271	4.808555	0.3075	4
4	7.644779	0.4689	7.997505	0.4337	8
5	9.093467	0.6949	9.554304	0.6550	12
6	20.42870	0.2015	21.92002	0.1458	16
7	24.17146	0.2350	26.06584	0.1636	20
8	24.59157	0.4282	26.53846	0.3264	24

*The test is valid only for lags larger than the VAR lag order.
df is degrees of freedom for (approximate) chi-square distribution
*df and Prob. may not be valid for models with exogenous variables

6.9 Residual Covariance and Correlation Matrices

	Exchange Rate	Czech CPI
Exchange Rate	1.164555	-0.008757
Czech CPI	-0.008757	0.050829

We observe just a small covariance of the residuals of the exchange rate and the Czech CPI.

	Exchange Rate	Czech CPI
Exchange Rate	1.000000	-0.035992
Czech CPI	-0.035992	1.000000

Their correlation is insignificant too, which is in no way surprising, as correlation is just a “normalized” version of covariance. Given tables inform us, that the orthogonal Cholesky impulses and just ordinary ones with the covariance matrix left unchanged might be quite similar.

6.10 Impulse Responses of CPI to One SD Shock of EX in the Table Representation

Period	1	2	3	4	5	6	7
	-0.008115	-0.008898	-0.048206	-0.054989	-0.071559	-0.077145	-0.082363

Corresponding descriptions, please, find in the main part of the paper.

6.11 Vector Autoregression Estimates

Vector Autoregression Estimates
Date: 12/21/12 Time: 22:49
Sample: 1999M01 2004M12
Included observations: 72
Standard errors in ( ) & t-statistics in [ ]

	D(EX_SA)	D(CPI_SA)
D(EX_SA(-1))	-0.045792	0.000848
	(0.10796)	(0.02256)
	[-0.42415]	[0.03760]
D(EX_SA(-2))	0.109811	-0.034078
	(0.11092)	(0.02317)
	[0.99002]	[-1.47061]
D(CPI_SA(-1))	-0.393321	0.209359
	(0.54827)	(0.11454)
	[-0.71739]	[1.82777]
D(CPI_SA(-2))	0.908055	0.287071
	(0.54718)	(0.11432)
	[1.65952]	[2.51121]
C	0.209218	0.023365
	(0.26467)	(0.05529)
	[0.79048]	[0.42255]
D(OIL)	-0.036219	0.012328
	(0.05754)	(0.01202)
	[-0.62945]	[1.02553]
D(CPI_US_SA(-1))	-0.182502	0.095007
	(0.41160)	(0.08599)
	[-0.44340]	[1.10486]
D(CPI_US_SA)	0.086840	0.127108
	(0.42361)	(0.08850)
	[0.20500]	[1.43626]
D3	-5.071532	0.087841
	(1.14068)	(0.23831)
	[-4.44607]	[0.36860]

D(OIL(-1))	0.009010	-0.024184
	(0.05548)	(0.01159)
	[0.16241]	[-2.08650]
R-squared	0.310584	0.289171
Adj. R-squared	0.210508	0.185986
Sum sq. resids	72.20243	3.151414
S.E. equation	1.079146	0.225453
F-statistic	3.103470	2.802459
Log likelihood	-102.2646	10.47376
Akaike AIC	3.118462	-0.013160
Schwarz SC	3.434666	0.303044
Mean dependent	0.198595	0.190069
S.D. dependent	1.214524	0.249886

Determinant resid covariance (dof adj.)	0.059117
Determinant resid covariance	0.043836
Log likelihood	-91.74422
Akaike information criterion	3.104006
Schwarz criterion	3.736413

6.12 Unit Root Tests for the Second Model

Variable	ADF		PP		KPSS	
	Level	1.Diff	Level	1.Diff	Level	1.Diff
ipi_sa	-1.595985	-5.137331	-1.892675	-5.044852	0.287725	0.413134
ppi_sa	-2.227330	-3.063912	-1.072594	-4.681533	0.138966	0.201092

6.13 Normality Test for the Second Model

VAR Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: residuals are multivariate normal
Date: 12/28/12 Time: 19:11
Sample: 1999M01 2004M12
Included observations: 72

Component	Skewness	Chi-sq	df	Prob.
1	0.015250	0.002791	1	0.9579
2	0.276070	0.914576	1	0.3389
3	0.167782	0.337808	1	0.5611
4	0.093486	0.104875	1	0.7461
Joint		1.360049	4	0.8511

Component	Kurtosis	Chi-sq	df	Prob.
1	4.287391	4.972130	1	0.0258
2	3.167210	0.083877	1	0.7721
3	2.738629	0.204944	1	0.6508
4	3.060930	0.011138	1	0.9160
Joint		5.272088	4	0.2605

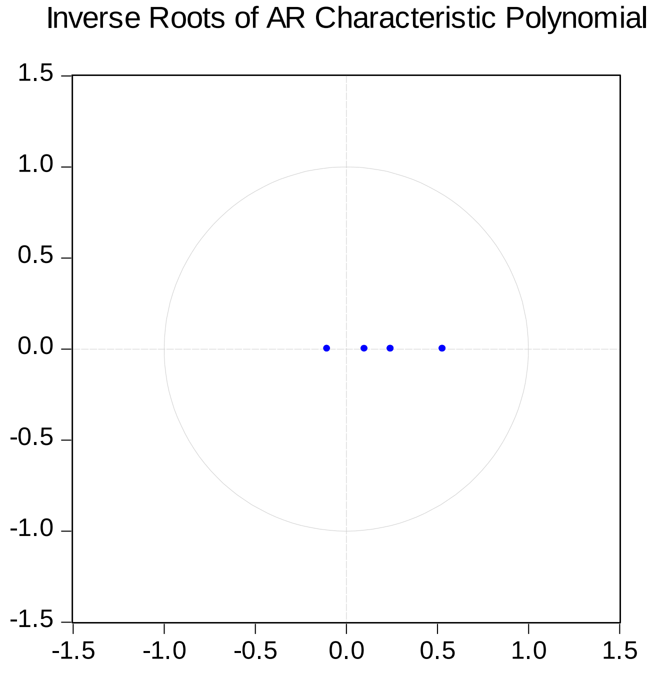
Component	Jarque-Bera	df	Prob.
1	4.974920	2	0.0831
2	0.998453	2	0.6070
3	0.542752	2	0.7623
4	0.116012	2	0.9436
Joint	6.632137	8	0.5768

6.14 Portmanteau Test for the Autocorrelations for the Second Model

VAR Residual Portmanteau Tests for Autocorrelations
Null Hypothesis: no residual autocorrelations up to lag h
Date: 12/28/12 Time: 19:12
Sample: 1999M01 2004M12
Included observations: 72

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	2.486751	NA*	2.521776	NA*	NA*
2	17.51881	0.9535	17.98332	0.9446	29
3	27.05438	0.9843	27.93348	0.9784	45
4	47.30401	0.9007	49.37427	0.8569	61
5	63.68020	0.8617	66.97256	0.7857	77
6	71.84354	0.9492	75.87802	0.9019	93
7	86.99476	0.9404	92.66091	0.8690	109
8	106.9433	0.8769	115.1030	0.7260	125

6.15 Inverse Roots of the Characteristic Polynomial of the Second Model



6.16 Residual Covariance and Correlation Matrices for the Second Model

	D(EX_SA)	D(IPI_SA)	D(PPI_SA)	D(CPI_SA)
D(EX_SA)	0.727047	-0.466130	-0.007336	-0.005957
D(IPI_SA)	-0.466130	0.719115	0.007851	0.027397
D(PPI_SA)	-0.007336	0.007851	0.065400	0.017793
D(CPI_SA)	-0.005957	0.027397	0.017793	0.055543

	D(EX _SA)	D(IPI_SA)	D(PPI_SA)	D(CPI_SA)
D(EX_SA)	1.000000	-0.644654	-0.033642	-0.029645
D(IPI_SA)	-0.644654	1.000000	0.036200	0.137087
D(PPI_SA)	-0.033642	0.036200	1.000000	0.295214
D(CPI_SA)	-0.029645	0.137087	0.295214	1.000000

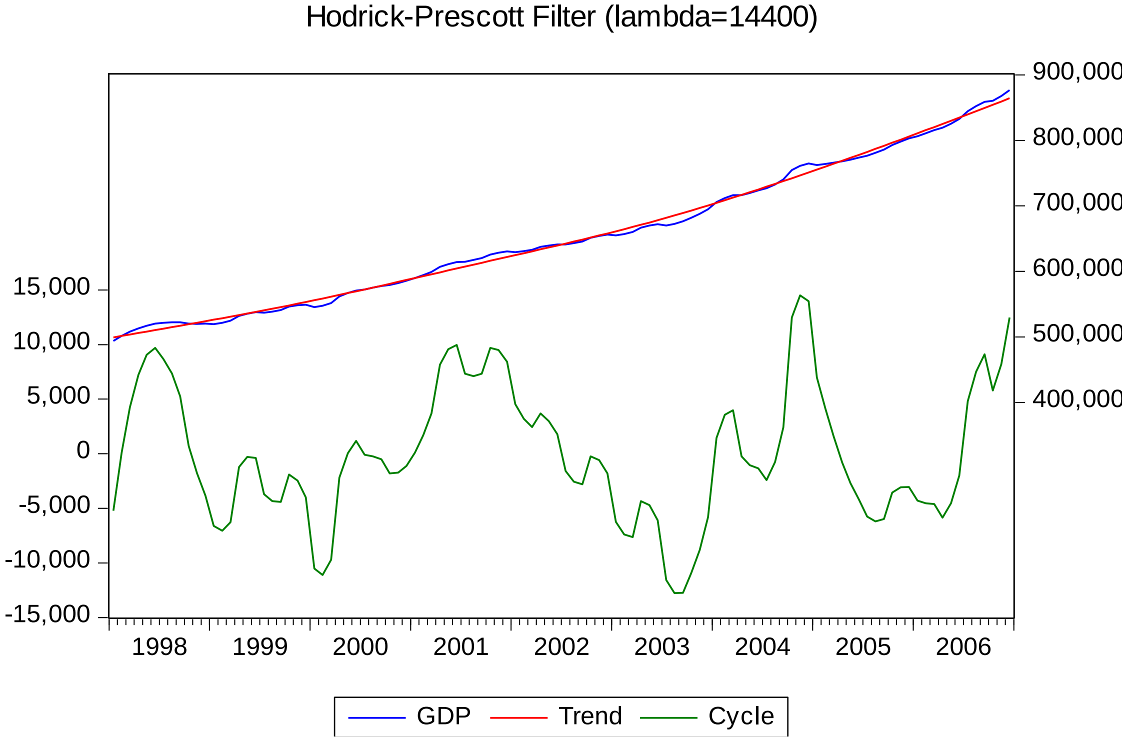
6.17 Vector Autoregression Estimates for the Second Model

	D(EX_LONG_SA)	D(IPI_SA)	D(PPI_SA)	D(CPI_SA)
D(EX_LONG_SA(-1))	-0.067690	-0.083115	0.037044	0.046033
	(0.13773)	(0.13698)	(0.04131)	(0.03807)
	[-0.49147]	[-0.60678]	[0.89678]	[1.20924]
D(IPI_SA(-1))	-0.102488	0.243463	0.105711	0.078606
	(0.15674)	(0.15589)	(0.04701)	(0.04332)
	[-0.65386]	[1.56179]	[2.24864]	[1.81441]
D(PPI_SA(-1))	0.428275	0.168737	0.449632	0.122277
	(0.34421)	(0.34232)	(0.10323)	(0.09514)
	[1.24424]	[0.49292]	[4.35545]	[1.28527]
D(CPI_SA(-1))	-0.239412	0.060656	0.010386	0.149349
	(0.46231)	(0.45978)	(0.13866)	(0.12778)
	[-0.51786]	[0.13192]	[0.07491]	[1.16878]
C	0.252048	-0.060459	0.080286	0.130921
	(0.14016)	(0.13939)	(0.04204)	(0.03874)
	[1.79829]	[-0.43373]	[1.90989]	[3.37951]
D(OIL_SA)	0.040787	0.140780	0.039496	0.017975
	(0.04159)	(0.04136)	(0.01247)	(0.01150)
	[0.98065]	[3.40340]	[3.16619]	[1.56359]
GAP	-1.97E-05	-1.46E-05	3.08E-06	1.88E-05
	(3.9E-05)	(3.8E-05)	(1.2E-05)	(1.1E-05)
	[-0.51024]	[-0.38176]	[0.26676]	[1.76803]
D(OIL_SA(-1))	-0.003847	-0.023501	0.007565	-0.034367
	(0.04947)	(0.04920)	(0.01484)	(0.01367)
	[-0.07776]	[-0.47766]	[0.50989]	[-2.51337]
GAP(-1)	5.84E-05	-3.03E-05	-2.43E-06	-1.20E-05
	(4.0E-05)	(4.0E-05)	(1.2E-05)	(1.1E-05)
	[1.45784]	[-0.76102]	[-0.20265]	[-1.08755]
D8	-3.932181	1.079596	-0.208776	-0.038853
	(0.88741)	(0.88256)	(0.26615)	(0.24528)

	[-4.43107]	[1.22326]	[-0.78442]	[-0.15840]
R-squared	0.338399	0.407941	0.500162	0.248717
Adj. R-squared	0.242361	0.321997	0.427604	0.139660
Sum sq. resids	45.07690	44.58512	4.054781	3.443636
S.E. equation	0.852670	0.848006	0.255734	0.235675
F-statistic	3.523570	4.746589	6.893344	2.280613
Log likelihood	-85.30491	-84.91000	1.400124	7.281398
Akaike AIC	2.647359	2.636389	0.238885	0.075517
Schwarz SC	2.963562	2.952593	0.555089	0.391720
Mean dependent	0.200232	0.098538	0.213760	0.191262
S.D. dependent	0.979603	1.029872	0.338018	0.254084

6.18 Construction of GAP Series

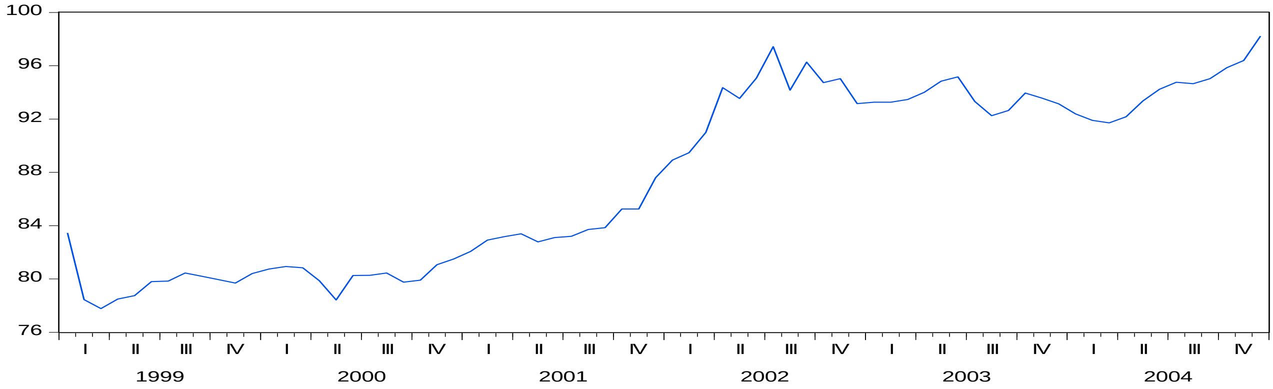
In order to construct monthly gap series, we apply Hodrick-Prescott filter on the monthly GDP observations:



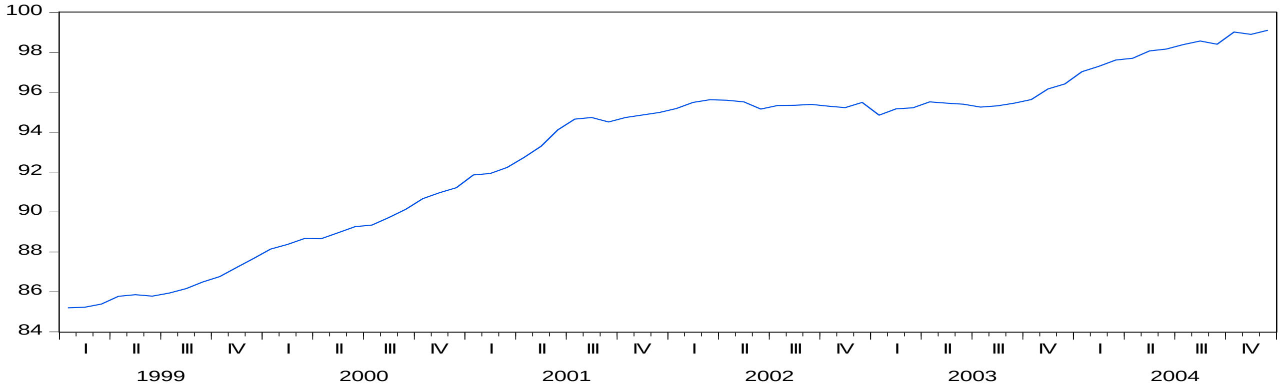
where GAP series is represented by the "Cycle" series.

7 Descriptive statistics

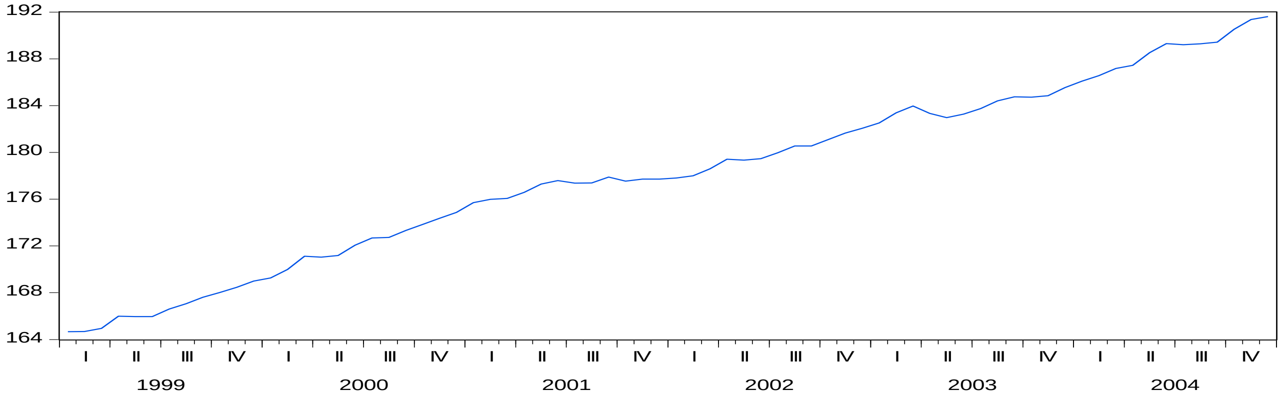
Nominal Effective Exchange Rate



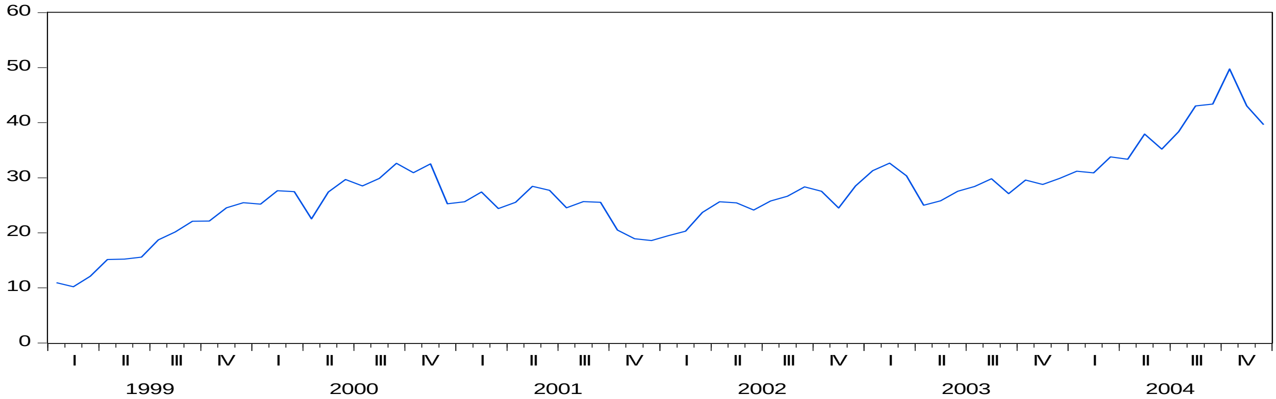
Domestic CPI



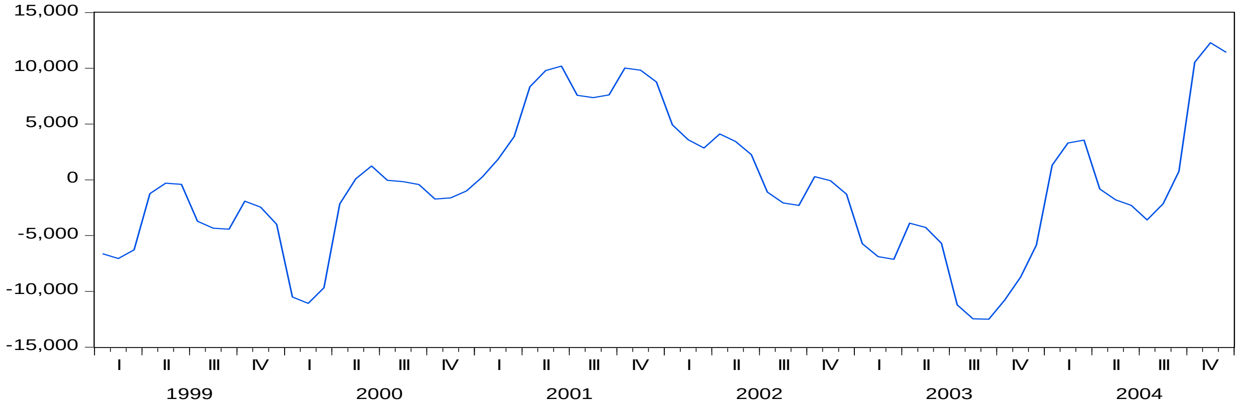
USA CPI



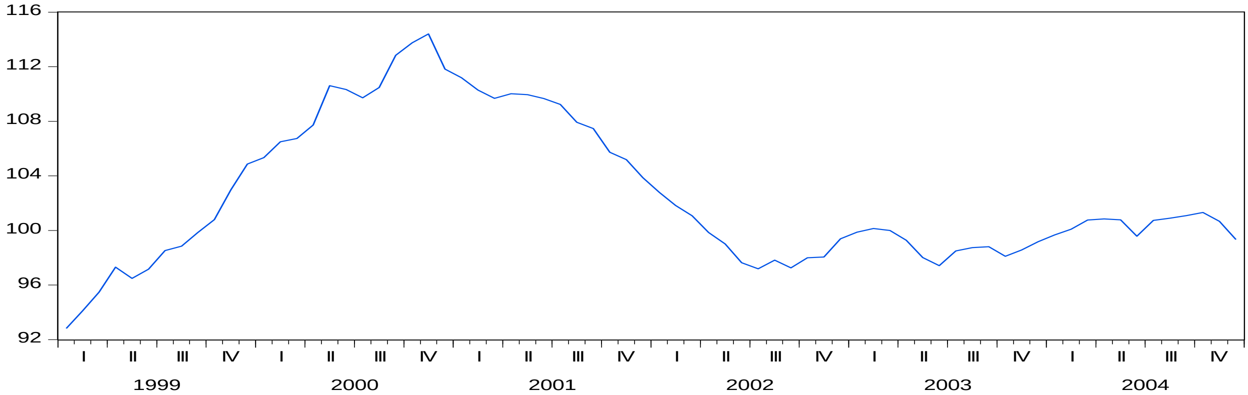
OIL



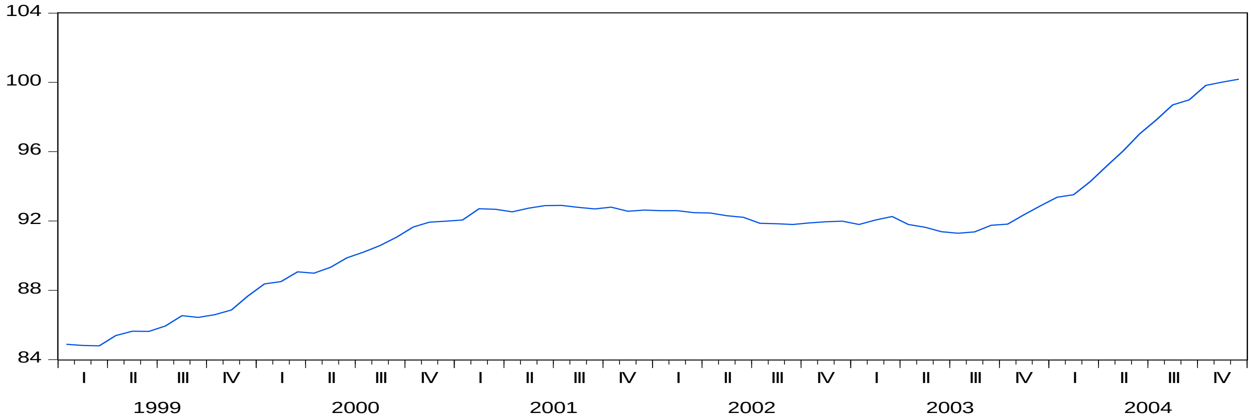
GAP



IPI_SA



PPI_SA



	CPI_SA	CPI_US_SA	EX_SA	OIL
Mean	93.07037	178.0947	87.55795	27.09444
Median	94.92486	177.8546	88.25382	27.25000
Maximum	99.10863	191.6205	98.22057	49.77000
Minimum	85.19986	164.6700	77.77921	10.20000
Std. Dev.	4.157125	7.499625	6.611837	7.309302
Skewness	-0.547360	-0.109836	-0.028691	0.356531
Kurtosis	2.030466	2.054522	1.310734	4.075805
Observations	72	72	72	72

	GAP	IPI_SA	PPI_SA
Mean	-587.4970	102.5259	91.63793
Median	-1049.529	100.7500	91.97049
Maximum	12281.10	114.3975	100.1847
Minimum	-12490.46	92.81149	84.78753
Std. Dev.	6115.842	5.287508	3.548064
Skewness	0.173324	0.611354	0.182695
Kurtosis	2.556808	2.220570	3.417103
Observations	72	72	72