State space model for handling VAT revisions

**Keywords:** State space model, incomplete data set, administrative data source, flash estimation

# Introduction

Aim of this research was to investigate whether an incomplete VAT-declaration dataset can be used for GDP flash estimation. The uncertainty around the early quarterly growth rates based on VAT data is relatively large. It can occur that some subjects to VAT do not provide tax return until the deadline, and furthermore, they have possibility to revise their tax returns after the first submission. The dataset could be finalised after six months after the reference year. Although VAT data do not conceptually identical to national accounts concept, it has the great advantage of being available early in contrast the survey data. But the revision of VAT data should be handled to be suitable for GDP flash estimation.

The dataset for the analysis contains monthly Hungarian VAT data for all data submission from 2015-2020 at enterprise level. Fiscal enterprises units do differ from statistical enterprise units as derived from the business register. In order to dealing with this problem, similar to the concept of Vlag and van den Bergen (2010), the analyse was not performed according to NACE Rev. 2,, but for high aggregation levels, for the total economy. Foreign enterprises without production in Hungary do not belong the enterprise population and were excluded from the VAT data. The most enterprises declare VAT on a monthly or quarterly base. Extremely small enterprises may declare their VAT on a yearly base. However, taking into their low contribution to total output, their VAT data can be neglected. Using historical VAT data of all available VAT submission from the tax office to Hungarian Central Statistical Office, a vintage matrix is compiled in order to forecast the revision of VAT data.

# Methods

There is used a state space model for forecasting of VAT revisions using the model of Cunningham et al (2007). During the data preparation phases the outliers are detected and handled. Then, as first step is compiled a vintage matrix from the data submissions according to the Table 1.

Table 1. Vintage matrix



Source: our compilation

The columns of the above table contain the reference month, and the rows include the each data submissions. Using the vintage matrix the revisions for all data submissions are calculated. The revision matrix (W) contains VAT data revision from January 2015 to December 2019. The rows of the matrix contain the data revision relating to the last data submission.

The applied model has two equations. The following state equation (1) is described for the estimation:

(1)

where yt is the final VAT data for reference month t, which cannot be observed. According to the state equation the final VAT data is a q ordered autoregressive process with expected value and expected error .

The measurement equation has be written by the formula (2):

(2)

where

is the submitted VAT data in T month for the reference month t. shows the revision bias and the measurement error .

# Results

The state space model is fitted and the revisions are forecasted. The model estimates the VAT revisions for the total economy. The growth rates calculated in accordance with VAT revisions are suitable for GDP estimation.

# Conclusions

This study presents a method to forecast VAT revision. The incomplete VAT-dataset consists of all monthly and quarterly VAT-declaration up to 27 days after the end of the corresponding month. To estimate VAT turnover growth rates, the expected revisions should be taken into account. Our model is able to forecast VAT revisions and so it provides better quality of growth rates causing smaller revisions in GDP volume index.

# References

1. Vlag, P., van den Bergen, D. (2010) : The use of VAT for short term statistics: some quality aspects <https://www.ine.pt/filme_inst/essnet/papers/Session1/Paper1.3.pdf>
2. Cunningham, A., Eklund, J., Jeffery, C., Kapetanios, G. és Labhard, V. (2007): A state space approach to extracting the signal from uncertain data. Bank of England Working Paper, No. 336.