Group Assignment II
Advanced Econometrics

Emma Arussi, 2783830, e.s.arussi@student.vu.nl
Marta Chejduk, 2753970, m.m.chejduk@student.vu.nl
Niels van Herk, 2832099, n.a.a.van.herk@student.vu.nl
Tess Scholtus, 2831775, t.p.scholtus@student.vu.nl

Academic Year: 2024-2025

QUESTION 1.

Model description

The **AR(1)** model is represented as follows:

$$x_t = \mu + \delta x_{t-1} + \epsilon_t$$

The **SESTAR** model introduces a smooth transition based on x_{t-1}:

$$x_t = \mu + \left(\delta + \frac{\gamma}{1 + e^{\alpha + \beta x_{t-1}}}\right) x_{t-1} + \epsilon_t$$

Similarly, the **STAR** model uses z_{t-1} (an exogenous variable) as the transition variable:

$$x_t = \mu + \left(\delta + \frac{\gamma}{1 + e^{\alpha + \beta z_{t-1}}}\right) x_{t-1} + \epsilon_t$$

(semi)parametric or (semi)non-parametric

A probability model is classified as **parametric** if the parameter space is finite dimensional, and nonparametric if the parameter space is infinite dimensional. A **semi-parametric** model can divide its parameter space into a finite-dimensional part and and infinite dimensional part. Finally, a semi-nonparametric indexes the parameter space by the sample size T. It is important also to mention that a probability model is only a probability model if the innovations are random variables, which is the case here since they are described as IID.

All three models have a finite parameter space. Where the largest possible parameter space is $\theta = (\mu, \delta, \gamma, \alpha, \beta)$, to which you can also add the standard deviation or the variance as an additional parameter to be estimated. The smallest parameter space, for AR(1) is defined as follows $\theta = (\mu, \delta)$.

In conclusion, the AR, STAR, and SESTAR models are parametric when the error terms are i.i.d. with $\mathbb{E}(\epsilon_t^2) < \infty$.

Model nesting

For both SESTAR and STAR it is easy to see that they reduce to the AR(1) model when $\gamma = 0$. Both SESTAR and STAR nest the AR(1) However, SESTAR does not nest STAR or the other way around, since there is no way to manipulate the parameters to change for the different dynamic variables used in the models.

QUESTION 2.

We calculated for the non-linear least-squares the total loss function with the given parameter vector $\theta = (\mu, \delta, \gamma, \alpha, \beta)$ for the AR(1) model, the SESTAR model, and the STAR model. The total loss L for each model is presented below:

- **AR(1):** $L = 22.92$
- **SESTAR:** $L = 50.46$
- **STAR:** $L = 48.52$

The parameter vector used for the evaluation is $\theta = (0, 0.1, 2, 0, 3)'$ for the SESTAR and STAR models, while for the AR(1) $\theta = (0, 0.1)'$

QUESTION 3.

Non-linear least squares is an optimization technique that can be used to build models that show non-linearity in the data or parameters. This is important to execute on SESTAR and STAR models since they are non-linear in parameter, and have intraceable estimators. Therefore, we need to minimize the errors by a minimization method, starting from an initialized vector for Theta. For this, appropriate bounds for the optimization of Theta have been set, and the BFGS method was used. The the results are displayed in Figure 1 below.

Figure 1 shows the models (AR(1), SESTAR and STAR in red, blue and green, respectively) fitted to the (differenced) unemployment rates. On the y-axis you see the unemployment rate and on the x-axis you see the date from 1959 until 2019.

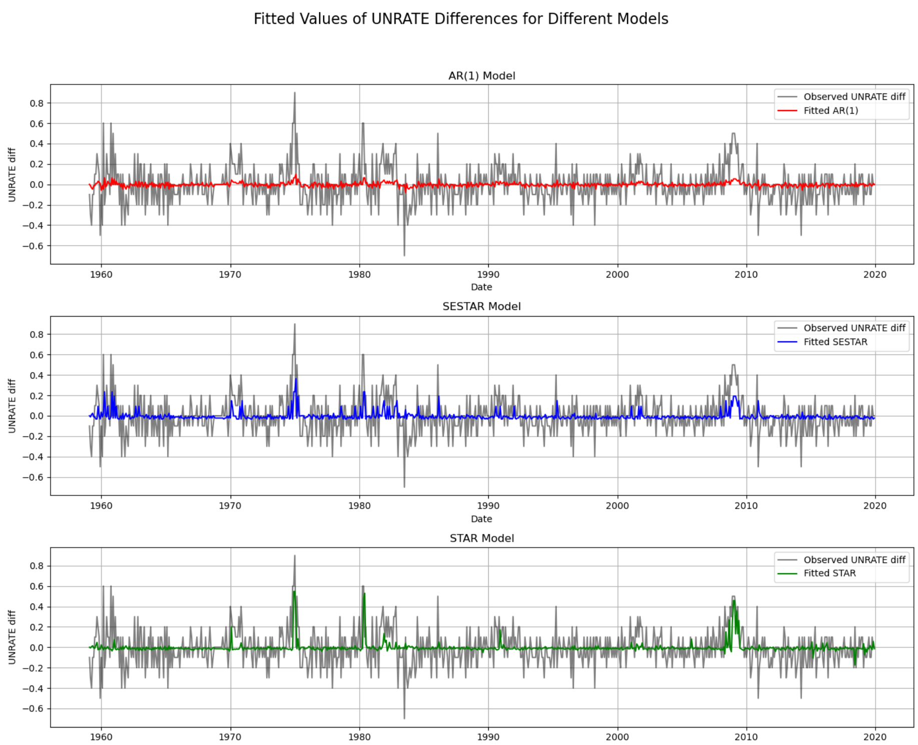

Figure 1: AR(1), SESTAR and STAR fitted on first difference of unemployment rates

We observe 3 periods of high unemployment rate values, the periods are: November 1973 - March 1975, January 1980 - July 1980 and December 2007-July 2009. During those periods there were recession (Oil crisis in 1973-1975, Recession 1980, Great recession

2007-2009) and therefore a higher unemployment rate.

Looking solely at the model fit from the impression of the plots and not the parameters or total loss, it seems the STAR model best predicts the periods of volatility.

QUESTION 4.

In figure 2 we observe the differenced unemployment rates between 1959 and 2019, as well as the logistic term of the star model, given as;

$\delta + \frac{\gamma}{1+e^{\alpha+\beta z_{t-1}}}$.

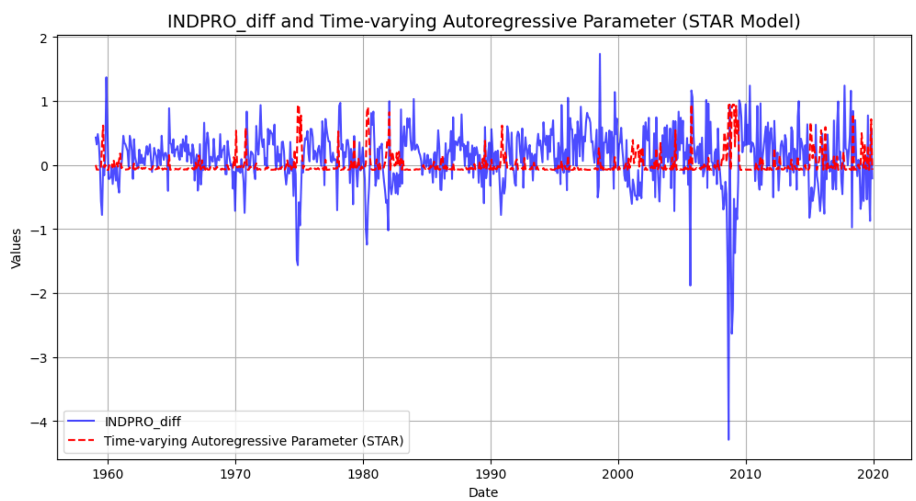

Figure 2: STAR model fit, as well as logistic term

The logistic term shows a negative relationship with the industrial production rates. This means, that if the industrial production rate is high, the star model reflects this as a negative effect on the unemployment rate. This seems like a logical effect, since industrial production in general affects jobs in real life as well.

QUESTION 5.

Description of the ACF plot

The Autocorrelation Function plot is presented in Figure 2.

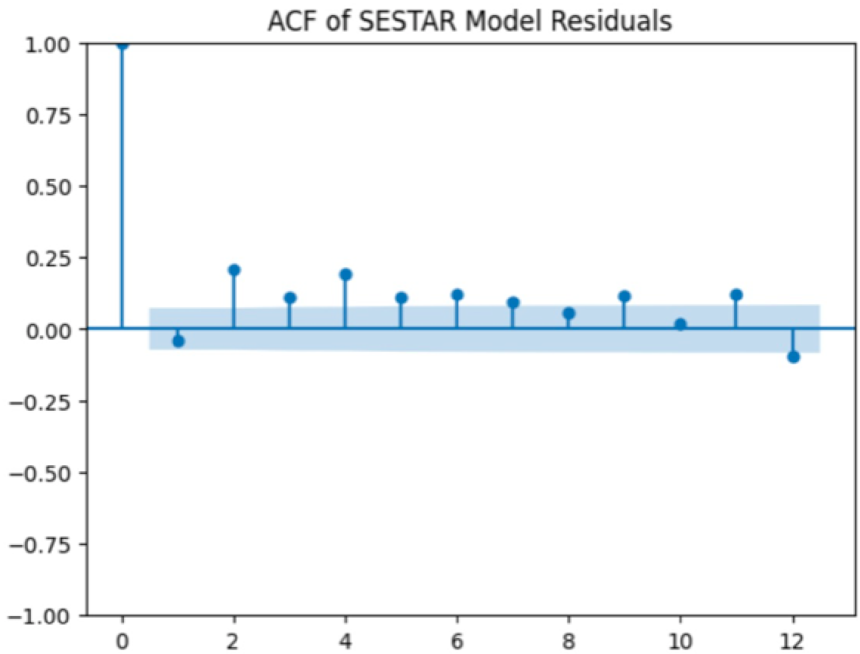

Figure 3: Estimated ACF of SESTAR model.

The plot shows the autocorrelations of the residuals of SESTAR model up to lag 12. The blue-shaded region represents the confidence interval. If a point lies outside this region, it indicates a statistically significant autocorrelation at that lag.

The autocorrelation persists even at lag 12, this is a strong indication for non stationarity, and therefore also misspecification of the model. To further research this, one could perform an augmented dicky fuller test to test stationarity. Furthermore, since the model could be misspecified, robust standard errors could increase model performance.

QUESTION 6.

Formulas used to calculate the robust standard Errors

The robust standard errors are calculated as the square root of the diagonal elements of $\hat{\Sigma}$.

First, we estimate

$$\Sigma := \operatorname{Var}\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\nabla q(x_t, x_{t-1}, \theta_0)\right)$$

with the Newey-West estimator

$$\hat{\Sigma} = \hat{\Sigma}_0 + \sum_{j=1}^{p}\left[1 - \frac{j}{p+1}\right]\left(\hat{\Sigma}_j + \hat{\Sigma}_j'\right),$$

where

$$\hat{\Sigma}_j = \frac{1}{T}\sum_{t=j+2}^{T}\nabla q\left((x_t, x_{t-1}, \hat{\theta})\right)\nabla q\left((x_{t-j}, x_{t-1-j}, \hat{\theta})\right)'.$$

AR(1) model

The scores of the non-linear least squares residuals with respect to the parameters μ and δ are presented below.

For μ:

$$\nabla_\mu q(x_t, x_{t-1}, \hat{\theta}) = -(x_t - (\mu + \delta x_{t-1}))$$

For δ:

$$\nabla_\delta q(x_t, x_{t-1}, \hat{\theta}) = -x_{t-1}(x_t - (\mu + \delta x_{t-1}))$$

The ∇q_t is the vector of the above formulas. Furthermore, the formulas for Σ_0 and Σ_j are the following.

$$\Sigma_0 = \frac{1}{T}\sum_{t=1}^{T}\nabla q_t \nabla q_t'$$

$$\Sigma_j = \frac{1}{T}\sum_{t=j+1}^{T}\nabla q_t \nabla q_{t-j}'$$

The above Σ_0 and Σ_j are used to calculate The Newey-West variance-covariance estimator. The variance estimates for parameters μ and δ are the diagonal elements of $\hat{\Sigma}$. This will be a 2x2 matrix.

$$\hat{\Sigma} = \begin{bmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} \end{bmatrix}$$

So,

$$\text{Var}(\mu) = \hat{\Sigma}_{11}, \quad \text{Var}(\delta) = \hat{\Sigma}_{22}$$

Furthermore, the robust standard errors of the parameters are the square roots of the diagonal elements of the Newey-West variance-covariance matrix.

$$\text{SE}(\mu) = \sqrt{\hat{\Sigma}_{11}} \quad \text{SE}(\delta) = \sqrt{\hat{\Sigma}_{22}}$$

STAR and SESTAR model

For STAR and SESTAR, the scores are calculated using jacobian() function from numdifftools package in Python. This function calculate derivatives numerically. The formula, which is a gradient of the the log-likelihood with respect to the parameter vector $\theta = [\mu, \delta, \gamma, \alpha, \beta]$ is given below;

$$S(\theta) = \frac{\partial \left(\sum_{t=1}^{T} (x_t - \hat{x}_t(\theta))^2 \right)}{\partial \theta}$$

Where the $\hat{x}_t(\theta)$ is the predicted value of x_t given the model parameters θ, which for SESTAR correspond to

$$\hat{x}_t(\theta) = \mu + \left(\delta + \frac{\gamma}{1 + e^{\alpha + \beta x_{t-1}}} \right) x_{t-1}$$

and for STAR to;

$$\hat{x}_t(\theta) = \mu + \left(\delta + \frac{\gamma}{1 + e^{\alpha + \beta z_{t-1}}} \right) x_{t-1}$$

Next, we calculate $\hat{\Sigma_0}$ and $\hat{\Sigma_j}$. At lag 0, the covariance matrix is:

$$\hat{\Sigma}_0 = \frac{1}{T} \sum_{t=2}^{T} q(x_t, x_{t-1}, \hat{\theta}) q(x_t, x_{t-1}, \hat{\theta})'$$

The lagged covariance matrix $\hat{\Sigma_j}$ where j = 1, 2... p is calculated the following way.

$$\hat{\Sigma}_j = \frac{1}{T}\sum_{t=2}^{T} q(x_t, x_{t-1}, \hat{\theta}) q(x_{t-j}, x_{t-1-j}, \hat{\theta})'$$

Now, calculate the Newey-West variance estimator is then a 5x5 matrix for both, since there are 5 parameters.

$$\hat{\Sigma} = \begin{bmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} & \hat{\Sigma}_{13} & \hat{\Sigma}_{14} & \hat{\Sigma}_{15} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} & \hat{\Sigma}_{23} & \hat{\Sigma}_{24} & \hat{\Sigma}_{25} \\ \hat{\Sigma}_{31} & \hat{\Sigma}_{32} & \hat{\Sigma}_{33} & \hat{\Sigma}_{34} & \hat{\Sigma}_{35} \\ \hat{\Sigma}_{41} & \hat{\Sigma}_{42} & \hat{\Sigma}_{43} & \hat{\Sigma}_{44} & \hat{\Sigma}_{45} \\ \hat{\Sigma}_{51} & \hat{\Sigma}_{52} & \hat{\Sigma}_{53} & \hat{\Sigma}_{54} & \hat{\Sigma}_{55} \end{bmatrix}$$

So,

$$\mathrm{Var}(\mu) = \hat{\Sigma}_{11}, \quad \mathrm{Var}(\delta) = \hat{\Sigma}_{22} \quad \mathrm{Var}(\gamma) = \hat{\Sigma}_{33} \quad \mathrm{Var}(\alpha) = \hat{\Sigma}_{44} \quad \mathrm{Var}(\beta) = \hat{\Sigma}_{55}$$

Finally, the robust standard errors are the square roots of the diagonal elements of $\hat{\Sigma}$.

The Results

The Results for three models are presented in the Table 1.

Table 1 shows the parameter estimates for AR, SESTAR, STAR models. Furthermore, for each model loss function, R^2 and R^2_{adj} are presented.

	AR	**SESTAR**	**STAR**
μ	-0.003	-0.022	-0.014
	(0.540)	(0.507)	(0.494)
δ	0.112	0.429	-0.075
	(0.131)	(0.101)	(0.081)
γ		-0.546	1.020
		(0.054)	(0.022)
α		-6.144	3.367
		(0.003)	(0.008)
β		28.518	5.277
		(0.001)	(0.004)
Loss	22.913	21.760	20.824
R^2	0.012	0.062	0.103
R^2_{adj}	0.011	0.057	0.098

QUESTION 7.

To formally test whether there are non-linear time-dependence effects in the SESTAR and STAR model a t-test for non-linear time dependence effects is used. This t-test is performed over the γ in the SESTAR and STAR model, since in both models setting $\gamma = 0$ implies that there are no non-linear time dependence effects. The t-statistics in the t-tests for the SESTAR and STAR model are calculated in the following way:

$$t_{\text{SESTAR}} = \frac{\hat{\gamma}_{\text{SESTAR}}}{\text{SE}_{\text{robust}}(\hat{\gamma}_{\text{SESTAR}})}$$
$$t_{\text{STAR}} = \frac{\hat{\gamma}_{\text{STAR}}}{\text{SE}_{\text{robust}}(\hat{\gamma}_{\text{STAR}})}$$

Next, the corresponding p-value to each of the t-statistics is used for the following hypothesis test:

$$H_0 : \gamma = 0 \quad \text{(no non-linear time-dependence effects)}$$
$$H_1 : \gamma \neq 0 \quad \text{(significant non-linear time-dependence effects)}$$

Table 1. The results of a t-test for non-linear time dependence effects in the SESTAR and STAR model.

The table shows the t-statistics and corresponding p-values of a t-test on the parameter γ in the SESTAR and STAR model.

Model	**t-statistic**	**p-value**
SESTAR	-10.1265	0.0
STAR	45.7071	0.0

In table 2 can be seen that for both the SESTAR and STAR model the $p_{\text{value}} < 0.05$ indicating that we reject H_0. Hence, from the results of the t-test performed on γ in the SESTAR and STAR model can be concluded that there are non-linear time-dependence effects in the SESTAR and STAR model.

QUESTION 8.

As the results in question 7 indicates that there are non-linear time-dependence effects in the SESTAR and STAR model, it is expected that these models fit the data better than the linear AR1 model.

Subsequently, in table 1 can be seen that the SESTAR and STAR models have a significantly larger R^2_{adj} then the linear AR1 model. Since the R^2_{adj} includes a penalization for the complexity of the model, this statistic is more appropriate to compare the fit of models with different complexity then using the regular R^2. Furthermore, the STAR model has a larger R^2_{adj} then both the SESTAR and AR1 model, hence it can be concluded that the STAR model fits the data the best.