The aim of the current study is to explore how the carbon tax influences the CO2 emissions in Sweden, during the period 1985-2021

Since the study focuses on the examination of one to one causal relationship among per capita carbon emission and other factors, it is fundamental to choose the appropriate specification of the model. Some of them include the vector error correction model (VECM), the vector autoregressive model (VAR) and the autoregressive distributed lag model (ARDL). The selection of these models depends on the unit root of the variable and on the existence of co-integration between the considered variables.

A VECM is similar to the VAR model but with the addition of stationary growth variables and an “error correcting equations with non-stationary level variables” (Lee, 2022).

The justification for VECM specification is based on the following. First, it has a beautiful interpretation using both long- and short-term equations to examine the impact of carbon tax rate on carbon emissions. Second, VECM is one of the time series modelling techniques that can directly predict the level at which a variable can be returned to equilibrium following a shock to other variables (Usman et al., 2017). Third, the VECM model is a specialized version of vector autoregression that works with variables that are integrated of order one, co-integrated, and have a long-run relationship. Fourth, VECM is employed because it deals appropriately with the order of integration of the selected variables based on the unit root test result. Consequently, a long-run connection between the series is required for VECM.

The equations for the VECM can be written as follows:

In the equations above,

represents the coefficient for the error correction term (hence the acronym ETC); , , , , are the constant terms; ∆ is the first difference of all the variables included in the model; , , , and are the short-run coefficient in the long-run equilibrium; are the residuals in the equations; k is the optimal lag length.

In order to develop the VECM, it is necessary to perform some step that are illustrated in the following chapters. These are the unit root test, the test for lag selection and the test for cointegration

## **2.0**

## **Econometric specification**

The following paragraph describes the econometrical specifications and the undertaken procedures of this study. As said before, the method used was time series, and the baseline model specification is represented in the equation 2 below.

  (2)

In this equation, α represents the intercept and βs are representing the coefficient values of independent variables, while  represents the error term.

As above-mentioned, it is necessary to transform the econometric model above in log form. Therefore, the equation is as follows:

     (3)

In equation 3, “ln” stands for the natural log of the respective variables.

3

Unit root test

As I am dealing with a time-series analysis, the first step to undertake is the unit root test. This test is fundamental, as the choice of the model directly depends on the presence or absence of unit root. The reason why is crucial to check the stationary is due to the fact that if variables are not stationary, there is the risk of a spurious regression (Baumohl and Lyocsa, 2009).

In time series, with the term stationary, we mean the situation in which “mean and variance are constant over time” (Pulina, n.d.). The specific test used is the Augmented Dickey-Fuller (ADF), as it is by far the most common statistical approach used in the previous studies, but other unit root tests include the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test, the Phillips-Perron test and the ADF-GLS test.

The unit root exists in time series of the value of α-1 in the given equation:

                                                       (4)

where is the value of time series at time t and is an exogenous variable. Therefore, if α is equal to 1, it means that it exists unit root in the dataset. The Dickey-Fuller (DF) test verify the null hypothesis that α=1 in the following equation, where represents the first lag of time series,  represent the first difference of the time series at  period (Prabhakaran, 2022)

= c+βt+α (5)

The Augmented Dickey Fuller Test evolved based on the above DF test and expands the DF test to include higher order regressive process in the model (Gujarati, 2003).The ADF test allows to incorporate additional lag which takes into consideration the possibility of serial correlation in the series. The null hypothesis of unit root is tested successively for each series by estimating regression equations for a random walk with a drift and trend, then a random walk with drift and finally a random walk. The advantage of using ADF is that it incorporates additional lag which takes into consideration the possibility of serial correlation in the series.

 The ADF test equation is given as below:

     (6)

From equation 6, it can be seen that the equation 5 has been extended only by adding more differencing terms while rest of equation remained same, which adds more thoroughness to the test.

The Augmented Dickey-Fuller generalized test equation is expressed as the follows. The first equation, it is constant with trend. In the second equation, is constant without a trend, and in the last one, it’s neither constant nor with trend

In these equations above, α is the intercept, namely the drift; β is the coefficient; t is the time; y is “the coefficient presenting process root” (RTC Lab, 2020) and is the residual term.

In order to correctly calculating the ADF test, it is important to select the correct number of lags. To do that, I adopted the Akaike Information criterion (AIC), named after the Japanese statistician Hirotugu Akaike, for each of the variables individually.

The results from the ADF test can result in different outcomes. If the variables included in the model are stationary at level, a simple regression is needed. If the variables are stationary at level, after first difference or mixed, it will be necessary to adopt the ARDL model. Finally, if the variables are stationary at first different, it is necessary to proceed with either VECM or VAR.

In this study, the variables are stationary at first difference. This result has been verified by comparing the t statistic with the critical value. I have used one as a lagged difference.

As visible from the stata output below (table 1), the variables are not stationary at level, as the absolute value of the t statistic is lower than the critical value at 1, 5 and 10%. Therefore, we cannot reject the null hypothesis. For this reason, the next step is to verify if the variables are stationary at first difference (table 2).

Table 1Augmented Dickey-Fuller test in level

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Test statistic | 1% critical value | 5% critical value | 10% critical value | P value |
| logco2em | -1.926 | -4.352 | -3.588 | -3.233 | 0.6409 |
| logGDPpc | -1.073 | -3.730 | -2.992 | -2.626 | 0.7257 |
| logCarbonTax | -1.008 | -3.730 | -2.992 | -2.626 | 0.7502 |
| logCPI | -0.660 | -3.730 | -2.992 | -2.626 | 0.8569 |
| logenergy | -0.954 | -3.730 | -2.992 | -2.626 | 0.7698 |

Table 2 Dicket-Fuller test ar first differences

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Test statistic | 1% critical value | 5% critical value | 10% critical value | P value |
| logco2em | -3.799 | -3.736 | -2.994 | -2.628 | 0.0029 |
| logGDPpc | -5.574 | -3.736 | -2.994 | -2.628 | 0.0000 |
| logCarbonTax | -2.619 | -2.492 | -1.711 | -1.318 | 0.0075 |
| logCPI | -4.700 | -3.736 | -2.994 | -2.628 | 0.0001 |
| logenergy | -6.156 | -3.736 | -2.994 | -2.628 | 0.0000 |

### After carrying out the ADF tets, it is necessary to perform the lag length selection criterion and to perform a cointegration test, as it is necessary to establish a long-run relationship.

4

Lag length Selection Criterion

The unit root analysis confirmed that both linear regression and ARDL approaches are not suitable for this study, therefore, we need to focus on either the VAR or VECM.

Following the ADF test, it is necessary to compute the lag-order selection model, in order to quantify what is the optimal number of lags in the model. The existence of lags in an analysis is due to the fact that a change in one variable does not instantaneously change the other variable, but some time is required. In order to do this, it is necessary to perform the command “varsoc” in Stata. This procedure is crucial because a too-high number of lags would mean a higher standard errors for the coefficients, a loss of a degree of freedom and, in general, a more uncertain model; too low, on the other hand, would result in a mis specified model and in a loss of information.

Based on the previous literature, there is not a more prevalent number of lags; they range from one to 16 in Andersson (2019). Generally speaking, the number of parameters should be less than the number of data points.

There are different types of tests to verify the optimal number of lags, such as the Final Prediction Error (FPE), the Akaike Information Criterion (AIC), the Hannan-Quinn Information Criterion (HQIC), as well as Schwartz and Bayesian Information Criterion (SBIC). The latter is often considered by the literature as the most precise among all (Ozcicek and Douglas McMillin, 1999.

The table below provides a summary overview to the lag length selection criterion.

All these tests suggest that four is the optimal number of lags.

 Table 3: results of VAR selection order criteria

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Lag | LL | LR | df | P | FPE | AIC | HQIC | SBIC |
| 0 | 166.352 |  |  |  | 2.8e-12 | -12.4117 | -12.342 | -12.1697 |
| 1 | 277.87 | 223.04 | 25 | 0.000 | 3.8e-15 | -19.0669 | -18.6489 | -17.6153 |
| 2 | 320.595 | 85.449 | 25 | 0.000 | 1.2e-15 | -20.4303 | -19.664 | -17.769 |
| 3 | 373.17 | 105.15 | 25 | 0.000 | 3.1e-16 | -22.5515 | -21.4368 | -18.6804 |
| 4 | 484.561 | 222.78\* | 25 | 0.000 | 3.3e-18\* | -29.197\* | -27.7339\* | -24.1162\* |

Note: \* represents the optimal lag based on the specific criterion

5.0   
Johansen test for cointegration

(not included in this document, but the result is one cointegrating equation)

VECM estimation results

The short-run estimates from the VECM model are explained in the table 6 below.

It is possible to see the coefficient of the variables at the different lags, together with their standard errors and p values.

Table 6: short-run VECM Estimation results

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | Coeff. | Std. error | z | P>|z| |
| *D\_logco2em* | \_ce1 |  |  |  |  |
|  | L1. |  | .0303334 | -0.33 | 0.744 |
|  | logco2em |  |  |  |  |
|  | LD. | -.1906241 | .3631337 | -0.52 | 0.600 |
|  | L2D. | .1666742 | .3819651 | 0.44 | 0.663 |
|  | L3D. | -.0705369 | .4368066 | -0.16 | 0.872 |
|  | logGDPpc |  |  |  |  |
|  | LD. | -.2541111 | .3605075 | -0.70 | 0.481 |
|  | L2D. | -.2068782 | .3459274 | -0.60 | 0.550 |
|  | L3D. | .1783543 | .2737584 | 0.65 | 0.515 |
|  | logCarbonTax |  |  |  |  |
|  | LD. | .295783 | .1606057 | 1.84 | 0.066\* |
|  | L2D. | .0323735 | .1855107 | 0.17 | 0.861 |
|  | L3D. | -.1048598 | .152445 | -0.69 | 0.492 |
|  | logCPI |  |  |  |  |
|  | LD. | -.9793666 | 1.385701 | -0.71 | 0.480 |
|  | L2D. | -1.362357 | 1.38458 | -0.98 | 0.325 |
|  | L3D. | 2.408046 | 1.111148 | 2.17 | 0.030\*\* |
|  | logenergy |  |  |  |  |
|  | LD. | -.4961457 | .3307021 | -1.50 | 0.134 |
|  | L2D. | -.0852191 | .303902 | -0.28 | 0.779 |
|  | L3D. | .1593979 | .3700134 | 0.43 | 0.667 |
|  | \_cons | -.0344675 | .0443692 | -0.78 | 0.437 |

Note: \*\* five percent significant level, \* 10 percent significant level

The results from table 6 show that, in the short-term, logCPI at lag three is significant at 5%. This means that an increase of the logCPI at lag three of 1% leads to an increase of the logco2em of 2,40%. Similarly, the table indicates that, in the short-run, an increase of 1% of the log of the carbon tax at lag one increases the CO2 emissions by 0,29%, and this is significant at 10%.

On the other hand, the rest of the variables, either the ones that present a positive or a negative impact on the emissions, are not significant.

When it comes to the results in the long-run, the outcomes can be observed in table 7 below. Again, it is possible to observe the variables alongside their coefficient, standard errors and p values.

Table 7: long-run VECM Estimation results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Coefficient | Std. error | Z | P>|z| |
| logco2em | 1 |  |  |  |
| logGDPpc | -11.25805 | .3801176 | -29.62 | 0.000\*\*\* |
| logCarbonTax | | 2.277334 | .1021611 | 22.29 | 0.000\*\*\* |
| logCPI | 25.41988 | 1.72313 | 14.75 | 0.000\*\*\* |
| logenergy | -.4082331 | 1.268987 | -0.32 | 0.748 |
| \_cons | | -84.29586 |  |  |  |

Note: \*\*\* one percent significant level

First of all, what the results indicate are that the variables locco2em, logGDPpc, logCarbonTax and logCPI are significant at 1%. On the other hand, logenergy is not significant.

In order to correctly carry out the interpretation, it is crucial to remember that the sign of the coefficient has to be interpret in the opposite way, due to the method of normalization.

The following is the interpretation for each variable.

The results from table 7 shows that a 1% increase in the carbon tax is significantly correlated with a reduction in emissions by 2,27%. This is in line with the theoretical assumptions written before, that state that a carbon tax is a viable, feasible and effective way of reducing emissions, and the results show how beneficial a carbon tax may be for the environment. The decrease in emissions is not outstanding, as firms may continue to pollute as long as it is economically convenient, but nevertheless, it  still remains an efficient policy. In the long-run, in order to reach the objective of zero net emissions by 2050, the best approach is to constantly raising the carbon tax every year. This is an effective way to reduce emissions, and it also gives firms and businesses the right time to adjust and adapt to a carbon-free future.

It might be pointed out that the reduction in emissions is not remarkably high. This could be explained by four reasons. The first one is that the Swedish carbon tax rate is already the highest in the World, which means that the marginal reduction is lower than other carbon tax policies around the World. Secondly, even if the carbon tax would increase by a significant margin (for example tenfold), the emissions would not reach zero, as there are other drivers for emissions. Thirdly, a very high tax rate from a carbon levy does not necessarily mean that emissions could be reduced further. In fact, it is not implied that the revenues generated from the tax are fully redistributed in environmental governmental policies to fight global warming, and firms can keep polluting. Finally, several industries are exempted from the payment of the tax, and can still keep on polluting until new regulations will be introduced by policymakers.

On the other hand, table 7 shows that there is a positive relationship between the GDP per capita and the emissions. In fact, a 1% increase in the GDP per capita affects the CO2 emissions by a growth of 11,2%. This is a result that comes up against the theory behind the Environmental Kuznets Curve (EKC), named after the American economist Simon Kuznets. The theory suggests that environmental degradation increases when a Country is developing, but, as the GDP grows, the environmental levels improve over time (Perman and Stern, 2003). However, the EKC has been subjected to multiple studies that have criticized it to various degrees. Some scholars claim that it does not simply exist in the real World (Perman and Stern, 2003), while Cole, Rayner and Bates (1997) state that it is applicable only to local air pollutants.

The majority of scholars agree with the results from this analysis, i.e. the economic growth is a leading cause for CO2 emissions (Moyer, Woolley, Glotter and Weisbach, 2013). The fact that the growth in the GDP is responsible for global warming has been observed in every region of the Earth except for Latin America (Acheampong, 2018).

This does not mean that economic growth will always be an obstacle for reaching net neutrality. In fact, several economists suggest that, due to technological innovations, a shift from an industrial to a service-based economy and a slower population growth, it may be possible to reach the environmental objectives without sacrificing the economic development (Begum, Sohag, Abdullah and Jaafar, 2015).

Therefore, logGDPpec and logCarbon tax have asymmetric effects on the logco2em in the long-run.

In addition to this, an increase of 1% in the consumer price index is significantly correlated with a reduction in the emissions on a per capita basis by 25,4%. An explanation of this may be that when the level of prices rises, the population tends to limit consumption, hence reducing the emissions. This is the phenomenon that periodically occurs when there is an increase in oil or gas prices, in which households simply decide to drive less. For these reasons, several environmental economists have claimed that higher fuel prices are actually beneficial for the environment (Frankel, 2021).

The value of the coefficient is remarkably high, at 25.41. The explanation may rely on the fact that the value for log of the cpi is relatively low, as well as the difference between its maximum and minimum value. Consequentially, a slight increase of 1% generates notable changes in the emissions.

Finally, in the normalization report of Johansen, the variable logenergy is not significant, therefore, it does not contribute to the change of the emissions. There are different explanations for this finding. One of it can be attributed to the different structure of the energy consumption in Sweden, that over the years has constantly increased the share of electricity coming from non-fossil fuels. The share of energy, consumed in Sweden, coming from nuclear power and renewable energy increased from 55,2% in 1985 to 71,1% in 2020 (Ritchie, Roser and Rosado, 2022). The largest increase is represented by the energy consumption coming from wind turbines, which increased from virtually zero in 1991 to 11,3% in 2020; on the contrary, oil decreased from 31,1% of the total consumption to 24,7% of today.

IMPULSE RESPONSE FUNCTION

An impulse-response function, a common post-estimation of the VECM, describes the evolution of the variable of interest along a specified time horizon after a shock in a given moment (Alloza, n.d.)

The graphs of the impulse-response function (IRF) for each independent variable were added to this work, as it may be interesting to evaluate how one variable fluctuates due to a one-unit shock on the other one. The IRF helps to better understand the evolution of the variables in the model (Mohr, 2020).

The IRF can be seen lying withing the 95% confidence interval, showing that the results are statistically significant and lie withing the given standard deviation.

In the graph in the top-left, we can observe that a one-time shock in the carbon tax affects the emissions in a mixed way. In the short-run (the first three periods), the shock is positive. Later, it is reduced until it becomes negative after the 7th period.

In the graphs on the right, regarding the CPI and the energy consumption, it is easily visible a sharp decline of the orthogonalized IRF, followed by a quick bounce back to the previous level. This may indicate the presence of a specific structural shock that may be the result of a change in a policy.

The rest of the other graphs share a similar pattern: the emissions respond permanently negatively due to a shock of the independent variable, as the blue line is consistently always below 0.

**Table 8: graph of Impulse Response Function**



## 6

## Eigenvalue stability condition

In addition to the Jarque-Bera statistic, when dealing with a vector error-correction model (VECM), another test that should be undertaken is done through the command “vecstable”, that checks the eigenvalue stability condition in the model (StataCorp, n.d).

The following results shows the graphical version of the eigenvalue stability condition, and the conclusion is that the the VECM specification imposes 4 unit moduli.

**Figure 6: roots of the companion matrix**

