UNIVERSITY
OF OSLO

Master thesis

# The economic consequences of Covid-19 - A comparison between Norway and Sweden analyzed using macroeconometric models

Bjarte Vårvik

Master of Economic Theory and Econometrics
30 ECTS Credits

Department of Economics
Faculty of Social Sciences

May 2023

# Preface

This thesis marks the end of my five years at the University of Oslo. I am thankful for all that I have learned, and the people I have met along the way.

I would like to say a huge thank you to my supervisor, Ragnar Nymoen. Your knowledge and helpfulness have been incredible throughout the process, and the outcome would not have been the same without your input. Thank you also to friends and family for all your support.

All views are my own, and any mistakes or shortcomings are fully my responsibility.

# Abstract

In this thesis, the economic effects of Covid-19 and the policy responses in Norway and Sweden are researched. Because Norway and Sweden chose quite different policy responses to prevent the spread of the pandemic, it is of interest to compare estimates of covid effects on the two countries' GDP. Using a well-documented machine learning algorithm, I constructed multiple-equation aggregate models as well as final form equations for Mainland-Norway GDP and for Sweden. The models were simulated to define hypothetical counterfactual no-covid scenarios. The results indicate a statistically significant medium- to long-run covid effect on Mainland-Norway GDP. On the other hand, the results show no indication of a long-term effect on GDP in Sweden. Compared to pre-existing studies for Norway, the main impression is that the results are similar, but of slightly stronger magnitude. Both sample differences and differences in methodology can explain the differences. For Sweden, there have been surprisingly few studies of the numerical consequences of the pandemic (and the responses to it) on GDP. I found a stronger economic effect on GDP in Sweden than in the limitedly available comparable studies.

# Table of Contents

# List of tables



# List of figures

# Chapter 1: Introduction

## 1.1 Motivation

The COVID-19 pandemic has caused significant disruptions to the global economy since its emergence in late 2019. The outbreak has led to numerous policy responses, including restrictions on movement and economic activity, to curb the spread of the virus. These measures have had a significant impact on businesses and households, which has led to widespread job losses, reduced income, and an economic slowdown.

The pandemic and the measures to prevent its spread can be regarded as large exogenous shocks. It can also potentially provide useful knowledge of how to deal with similar shocks in the future. This makes it interesting to undertake research to study the effects empirically. However, the different effects are not easily distinguishable. To study the economic effect, model building that involve many simplifications are needed in order to reach conclusions that are based on analysis. One example of how to simplify the real world is by using a small, aggregate model of the economy, which will be the main research method of this thesis.

## 1.2 Economic consequences of Covid-19 in Norway and Sweden

Norway and Sweden are two countries that were significantly affected by the pandemic. Both countries implemented measures to contain the spread of the virus, including lockdowns, travel restrictions, and social distancing measures. However, the measures implemented in Norway where much stricter than in Sweden. Norway chose to shut down most of the country, including lockdowns, while Sweden chose a more lenient tactic in dealing with the virus. Because the countries are, in a broad perspective, otherwise similar, the difference in policy response can be a highly interesting research setting.

The pandemic and the policies have had significant economic consequences, and it is important to understand the extent of the impact to inform future policy decisions. A comparison of the economic impact of COVID-19 in Norway and Sweden could provide valuable insights into the economic impact of the pandemic and the effect of differences in policy regimes. Moreover, there have not been any measures in place to reduce the spread of the virus in either country since January 2022. This makes the time of writing (spring 2023) an ideal time to study the economic consequences ex post, when data on the period affected newly has become available.

Overall, a thesis on the economic impact of COVID-19 in Norway and Sweden would be a timely and relevant contribution to the field of economics. It would provide valuable insights into the impact of the pandemic on these countries, as well as improve the information on future policy decisions aimed at mitigating the economic fallout of similar crises. Did the difference in policy response lead to a difference in economic impact between the two countries? This question leads to the research question of this thesis:

**What were the economic effects of Covid-19 and the difference in policy responses in Norway and Sweden?**

The research question will be empirically analyzed using multiple-equation macroeconometric models of GDP as well as final form equations created using a machine learning algorithm. The analysis builds on the estimated hypothetical counterfactual economic development of the aggregate models assuming the pandemic had not occurred. In practice, this will be done using dynamic simulation, where the models will be forecasted using time series data up until the fourth quarter of 2019. The final form equations for GDP in Mainland-Norway and Sweden will yield a relevant comparable estimate for each country's development in GDP.

## 1.3 Outline

The second chapter presents other studies on the economic effects of Covid-19, as well as a paper with a notational framework that will be used in this thesis. The third chapter presents the variable definitions and the corresponding data, as well as several concepts from time series econometrics which leads up to the choice of method. Moreover, documentation of the machine learning algorithm that will be used in the modelling are presented in chapter three. In chapter four, re-estimations of existing macroeconometric models for Mainland-Norway and Sweden, and the modelling results based on the new dataset that I use, will be presented. Moreover, the chapter include estimation results for final form equations for the two countries' GDP. Results from the aggregate models and final form equations are presented in the fifth chapter. The results for Mainland-Norway and Sweden are then compared and discussed in light of the research question. The sixth and final chapter will conclude the thesis.

## 1.4 Data and software

For the modelling of the Norwegian economy, the data bank of the Norwegian Aggregate Model (NAM) (Bårdsen & Nymoen, 2022) is used. For the modelling of the Swedish economy, I have constructed a new data bank based on Macrobond (2023), and my own calculations that are documented separately in chapter 4.3 and 4.5.

The software used for modelling is the module PcGive 15 in OxMetrics 8 (Doornik & Hendry, 2018). Some data set manipulation and variable creation has been done in R 4.2.1 (R Core Team, 2022).

# Chapter 2: The literature on macroeconomic covid effects in Norway and Sweden

## 2.1 Reports written on behalf of the corona commissions in Norway and Sweden

At this point, there have been conducted a few studies on the economic effect of the Covid-19 pandemic on both the Norwegian and the Swedish economy, see Blytt et al. (2022), Andersen et al. (2022) and Bjertnæs et al. (2021). These studies have used economic forecasts from before the pandemic as a counterfactual, i.e., what would have happened in a no-covid scenario. The counterfactual has then been compared with a baseline combination of actual GDP numbers and forecasted GDP values towards the end of the forecast period. This is in line with the method that will be used in this thesis, except for the baseline solely being actual GDP numbers as this data is now available.

In a background report for the Swedish corona commission (Andersen et al., 2022), a comparison of the economic consequences of the pandemic in the Nordic countries is investigated. The report finds that the developments in all Nordic countries are relatively similar, but that there were some sectoral differences. Moreover, compared to most other countries in the world, the Nordic countries fare well, both in terms of number of cases and the impact on the economy. This corresponds well with another finding, namely that there is a strong correlation between performance in terms of health and economy.

Compared to the average decline in GDP globally (9.1%) and in the Euro area (14.6%), the decline was less severe for the Nordic countries. In 2020, the decline was 2-3 percent in Norway, Sweden and Denmark (Blytt et al., 2022). The report concludes that as per end of year 2021, the "economic consequences were similar between the Nordic countries, while the health consequences were more dire in Sweden". Bjertnæs et al. (2021) estimates a stronger decline in Mainland-Norway GDP. In this report, the decline in 2020 is estimated to 145 billion, equivalent to 4.7 percent. For the years 2020-2023, the total discounted loss in economic activity is estimated to 330 billion, or 11 percent of yearly GDP. Furthermore, the report looked at the effect of the financial stimulus, which is estimated to dampen the decline by 0.5 percentage points, but notes that this effect is likely to be somewhat underestimated.

It is common in the literature on Covid-19 to remind the reader that measuring the cost of the pandemic in GDP does not quantify the complete welfare loss. Only the loss in economic activity is measured and does not include the cost of loss in terms of health or death, reduced learning and education, nor is reduced life quality due to lockdowns and social distancing. This is an important distinction, and crucial to keep in mind when assessing the findings. However, despite its delimitations, using economic activity as measure captures important costs which can be a starting point for more thorough assessments.

The Scandinavian countries chose different strategies for dealing the coronavirus. Norway and Denmark chose quite strict measures from the beginning, while Sweden chose to be more lenient. This gives us something similar to a natural experiment (Blytt et al., 2022), which can shed light on the effect on economic activity between the countries. This does not account for changes in economic behavior between the countries. For example, a hypothesis could be that some of the difference in this setting could disappear due to higher behavioral adaption in Sweden relative to in Norway.

## 2.2 A macroeconometric study on Covid-19 effects in Norway

The approach that I follow in this thesis builds on Nymoen (2023), which looks at the economic effect of Covid-19 using macroeconometric models. Different from what has been done in the literature previously, I have used small, aggregate models of the economy, and extended the research question to include Sweden, for comparison purposes. The benefit of using a simpler modelling framework is that it is easier to implement and interpret the estimated coefficients. Furthermore, a larger model might be mis-specified, potentially leading to biased forecasts (Pesaran & Smith, 2016). On the other hand, it will rely more heavily on simplifications of the real world, and less of the historic variation will be explained by different explanatory variables.

Following is a recap of the framework presented in Nymoen (2023), where I have made some simplifications. First, a notational framework for a macroeconometric model consisting of a conditional and marginal model is presented. Thereafter, the concept of difference between a baseline and counterfactual solution, and the difference between actuals and counterfactual solution is presented.

A bivariate VAR(1) model can be re-written as:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \tag{2.1}$$

$$X_t = \phi_{20} + \phi_{21} T_{t-1} + \phi_{22} X_{t-1} + \varepsilon_t \tag{2.2}$$

The first equation for $Y_t$ is the conditional model, while the second equation for $X_t$ is the marginal model. This can be extended to n endogenous variables and m exogenous variables in period t in the following compact writing:

$$y_t = f_y\left(y_{t-1}, \dots, y_{t-p}, x_t, \dots, X_{t-p}, D_{yt}, \varepsilon_{yt}\right), \text{where } f_y(\bullet) \text{ denotes a function.} \tag{2.3}$$

As in Nymoen (2023), “$D_{yt}$ represents deterministic trends which may include constants, trends, seasonal dummies and indicator variables for interventions or shocks.” $\varepsilon_{yt}$ represents random error-terms. The error terms are, by conditioning on other arguments in the function, unpredictable by default.

The model can be extended by endogenizing the variables in the $x_t$ vector:

$$x_t = f_x\left(x_{t-1}, \dots, x_{t-p}, D_{xt}, \varepsilon_{xt}\right) \tag{2.4}$$

This extended model can be combined and written compactly as:

$$\boldsymbol{y}_t = f\left(\boldsymbol{y}_{t-1}, \dots, \boldsymbol{y}_{t-p}, \boldsymbol{D}_t, \boldsymbol{\varepsilon}_t\right) \tag{2.5}$$

In the extended model, $y_t$ and $x_t$ are stacked in a $m + n$ vector $\boldsymbol{y}_t$, $\varepsilon_{yt}$ and $\varepsilon_{yt}$ in $\boldsymbol{\varepsilon}_t$, and $D_{xt}$ and $D_{yt}$ in $\boldsymbol{D}_t$. It is a system of difference equations, with general functional form. Because the solution of $\boldsymbol{y}_t$ is a function of a set of initial conditions, it can be referred to as a causal solution. This differs from the forward solution which does not condition on initial conditions (Nymoen, 2019, p. 133)

The information set that the solution is based on is notated by $\mathcal{I}_D$ in the following. A solution based on $\mathcal{I}_{D=1}$ will be referred to as a baseline solution, while a solution based on $\mathcal{I}_{D=0}$ will be referred to as a counterfactual solution. The effects of a large shock in this setting can be defined as the difference between two conditional expectations:

$$Diff_I \boldsymbol{y}_t = E(\boldsymbol{y}_t^c \mid \mathcal{I}_{D=0} - E\big(\boldsymbol{y}_t^b \bigm| \mathcal{I}_{D=1}\big), t = 1,2,\dots,T \tag{2.6}$$

A different way to measure the economic effect of the pandemic, is to estimate the difference between economic forecasts from before the shock and the actuals under the pandemic. This way of measuring the consequences can be notated as:

$$Diff_{II} \boldsymbol{y}_t = E(\boldsymbol{y}_t^c \mid \mathcal{I}_{D=0,other}) - \boldsymbol{y}_t, t = 1,2,\dots,T \tag{2.7}$$

The expectation term above denotes a forecast estimated prior to the pandemic, while $\boldsymbol{y}_t$ denotes the actuals. Because the two information sets are different, the subscript "other" is used to differ between the two. More precisely, the forecast that is made before the pandemic, is typically made conditional on assumptions which is different from the conditional expectation in the first term. In addition, the second term in each equation differs. In $Diff_I \boldsymbol{y}_t$, the second term is the simulated conditional expectation, while in $Diff_{II} \boldsymbol{y}_t$ it is the actual. Following this, the two equations for the estimated difference between the baseline and the counterfactual are generally different.

It is the second measurement of difference, $Diff_{II} \boldsymbol{y}_t$ (2.7), that will primarily be used in this thesis. The definition of $Diff_I \boldsymbol{y}_t$ (2.6) is also included, as a combination of the two differences have been used in the research literature so far and will as such be a useful reference for later explanations.

# Chapter 3: Data and empirical methods

## 3.1 Data

To investigate the research question empirically, I have used one primary data source each for Norway and Sweden. For the empirical modelling of the Norwegian economy, Norwegian Aggregate Model (NAM) (https://www.normetrics.com/nam, 2022) was used. For the estimations of the Swedish economy, data have been collected by the use of Macrobond (2023). The license to Macrobond was provided to me by The Department of Economics at the University of Oslo. Macrobond is a software company that gives an interface to time series data from producers of statistics worldwide. In this case, the statistics collected from Macrobond were from Statistics Sweden (SCB), the Swedish National Financial Management Authority (Ekonomistyringsverket) and the European Central Bank (ECB).

For the Norwegian data, the sample period was from the first quarter of 1988 to the third quarter of 2022. The initial sample size was shortened somewhat due to the estimation period ending before the start of the pandemic, and due to including a lag of the variable with the shortest time series available. Consequently, the estimation sample used was 1988(2) – 2019(4), totaling to 127 observations. The last eleven observations, i.e., 2020(1) – 2022(3), were used as actuals, to compare the model-simulated GDP with actual GDP for these quarters.

For the Swedish data, the sample was from the first quarter of 1996 to the fourth quarter of 2022. Due to the same reasons as with the Norwegian data, the sample size was reduced by a few observations at the beginning of the sample (including lagged variables), and at the end of the sample (estimation period ending before the shock). The resulting estimation sample became 1996(3)-2019(4); 94 observations in total. While the Swedish data included the last quarter of 2022, it was excluded from the forecast evaluation, in order to match the forecasting lengths of the two equation systems.

All variables, except for the interest and inflation rates, were converted to logarithms before being used in the econometric equations. This means that in the main, the estimated coefficients are interpretable as elasticities. However, the coefficient of the real interest rate variable for instance, is a semi-elasticity.

## 3.2 Time series econometrics

In this thesis, the research question is tested empirically by constructing models for GDP using time series econometrics. This implies that the parameter estimation is done based on historical time series data. In the following, a series of principles in time series econometrics is presented. The presentation of the statistical methods in this chapter builds on chapters 3 and 5-10 in Nymoen (2019), and chapter 14 in Stock and Watson (2015).

### 3.2.1 Unit-root

A unit-root is referred to as when the characteristic root, also known as the eigenvalue, has a modulus equal to 1 (Nymoen, 2019, p. 119). In a homogenous difference equation with a single characteristic root, $\phi_1$, the equation is stable as long as $|\phi_1| < 1$. The opposite of a stable equation is an explosive one, where, given initial conditions, at least one of the characteristic roots are larger than 1. In the case when there is a unit-root, i.e., $\phi_1 = 1$, the solution is unstable.

### 3.2.2 AR-model

An autoregressive process of first order, AR(1), is an example of a stationary time series. The first-order dynamics of such a process can be defined by the following difference equation:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t,$$

$$where\ e_t \sim IID(0, \sigma_\varepsilon^2), \qquad t = 1, \dots, T$$

It follows that most of the process in period t is dependent on last period. The AR(1) equation has a globally asymptotically stable solution as long as the autoregressive parameter $\phi_1 \epsilon (-1, 1)$, i.e., the stability condition is satisfied. The autoregressive model equation for AR(1) can be extended to higher order AR dynamics, with p-lags. This gives use the general characterization of an autoregressive model: It estimates the current value of a variable using only the historic values of the variable (Stock & Watson, 2015, p. 580).

### 3.2.3 VAR

A vector autoregressive process is the vector extension of the AR models presented above. In its simplest form, the VAR equals the AR(1) model. However, a VAR generally includes more variables. More specifically, a vector of AR(p) models, p being the number of lags. It can be specified on the following matrix form:

$$\begin{bmatrix} Y_t \\ Xt \end{bmatrix} = \begin{bmatrix} \phi_{10} \\ \phi_{20} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{bmatrix}, \tag{3.1}$$

$$where \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{bmatrix} \sim IID \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} \right], \qquad t = 1, \dots, T$$

In the equation system above, $Y_t$ is the endogenous variable and $X_t$ the exogenous. The error terms are assumed to be independently and identically distributed (IID). The VAR can also be written using scalar notation:

$$Y_t = \phi_{10} + \phi_{11} Y_{t-1} + \phi_{12} X_{t-1} + \varepsilon_{yt} \tag{3.2}$$

$$X_t = \phi_{20} + \phi_{21} Y_{t-1} + \phi_{22} X_{t-1} + \varepsilon_{yt} \tag{3.3}$$

This corresponds to a conditional and marginal model. It can be rewritten as:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \quad (Conditional) \tag{3.4}$$

$$X_t = \phi_{21} Y_{t-1} + \phi_{22} X_{t-1} + \varepsilon_{xt} \qquad (Marginal) \tag{3.5}$$

The marginal model corresponds to the second row in the VAR model above. The conditional model equals an autoregressive distributed (ADL) model. The error term, $\varepsilon_t$, in the conditional model is Gaussian white noise:

$$\varepsilon_t = IID \sim (0, \sigma^2 | X_t, X_{t-1}, Y_{t-1}), \qquad t = 1, \dots, T$$

### 3.2.4 SEM

A simultaneous equation model (SEM) is a system of equations that are interrelated and must be solved together. In this type of model, each equation includes several variables, and the values of those variables are simultaneously determined by the other equations in the system. A SEM can help explaining complex relationships between economic variables. The key difference from a

VAR is that the equations in a SEM are explicitly specified and reflect a theoretical understanding of the relationship between the variables in the system.

### 3.2.5 Equilibrium correction model

The equilibrium correction model (ECM) is an autoregressive distributed lag (ADL) model which has been reparametrized (Nymoen, 2019, p. 238). I start by showing the reparameterization of an ADL to ECM, following the guidelines in Nymoen (2019).

An ADL model with one explanatory variable and one lagged variable can be written as:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \beta_0 X_{t-1} + \beta_1 X_{t-1} + \varepsilon_t, t = 1, \dots, T \tag{3.6}$$

$$where\ E(\varepsilon_t | Y_{t-1}, X_t, X_{t-1}) = 0$$

Using algebraic manipulation to transform the ADL model to an ECM:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t\ (ADL) \tag{3.7}$$

$$Y_t - Y_{t-1} = \phi_0 + \phi_1 Y_{t-1} - Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$$

$$\Delta Y_t = \phi_0 + (\phi_1 - 1)Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \beta_0 X_{t-1} - \beta_0 X_{t-1} + \varepsilon_t$$

$$\Delta Y_t = \phi_0 + \beta_0 \Delta X_t + (\phi_1 - 1)Y_{t-1} + (\beta_0 + \beta_1)X_{t-1} + \varepsilon_t\ (ECM) \tag{3.8}$$

When $-1 < \phi_1 < 1$, last period's deviation from equilibrium may be expressed as an explicit variable in the model equation. This means the above ECM-equation can be rewritten and transformed the following way:

$$\Delta Y_t = \beta_0 \Delta X_t + (\phi_1 - 1)\left\{Y_{t-1} - \frac{\phi_0}{(1 - \phi_1)} - \frac{\beta_0 + \beta_1}{1 - \phi_1} X_{t-1}\right\} + \varepsilon_t, \qquad \phi_1 < |1| \tag{3.9}$$

$$\Delta Y_t = \beta_0 \Delta X_t + (\phi_1 - 1)\{Y_{t-1} - \mu_{Y|X_{t-1}}\} + \varepsilon_t$$

$$where\ \mu_{Y|X_{t-1}} \equiv \frac{\phi_0}{(1 - \phi_1)} - \frac{\beta_0 + \beta_1}{1 - \phi_1} X_{t-1}$$

Letting $E(Y_t) = E(Y_{t-1}) = Y^*$ and $E(X_t) = E(X_{t-1}) = X^*$ denote a stationary solution, a long-run multiplier can be obtained.

$$Y^* = \mu_{Y|X_{t-1}} \equiv \frac{\phi_0}{(1 - \phi_1)} - \frac{\beta_0 + \beta_1}{1 - \phi_1} X_{t-1} \tag{3.10}$$

By taking the derivate of $Y^*$ with respect to $X^*$, the permanent effect on Y from a change in X can be denoted by:

$$\frac{\delta Y^*}{\delta X^*} = \frac{(\beta_0 + \beta_1)}{(1 - \phi_1)} \equiv \beta^* \tag{3.11}$$

### 3.2.6 Non-stationarity

The characterization of a stationary variable is that the unit-root associated with the homogenous part of a difference equation is less than one. A stationary variable will move towards a steady-state in the long-run. A non-stationary variable is the opposite, which will either increase or decrease infinitely. A stationary variable can be denoted as I(0), meaning that the order for integration is zero. Accordingly, a non-stationary variable is annotated as I(k), where k is the order of integration. Macroeconomic variables such as GDP tend to be non-stationary. In the case of non-stationarity, the model is a so-called random walk model:

$$Y_t = \phi_0 + Y_{t-1} + \varepsilon_t$$

The equation above is a random walk with a drift term when $\phi_0$ is non-zero. Because the variables in a random walk model are non-stationary, a non-standard distribution must be used when testing the parameters' significance. Using the standard statistical distribution will lead to spurious regression. A spurious regression is defined as two (or more) time series appearing to be related when they are in fact unrelated (Stock & Watson, 2015, p. 601).

### 3.2.7 Augmented Dickey-Fuller test

The Augmented Dickey-Fuller (ADF) test is used to determine whether a time series is stationary or not (Nymoen, 2019, p. 334). More concretely, it tests whether the mean and variance of the time series are constant over time. The test is an extension of the simpler Dickey-Fuller test, which assumes that the time series has a unit-root and test whether that unit-root exists.

The ADF-test can be written on the following form (Nymoen, 2019, p. 334):

$$\Delta Y_t = \phi_0 + \sum_{i=1}^{p-1} \phi_i^{\dagger} \Delta Y_{t-1} + \delta t - \pi(1)\mathrm{Y}_{t-1} + \varepsilon_t$$

Where $\Delta Y_t$ is the first difference of the time series $Y_t$, $\pi(1)$ is a coefficient representing the degree of persistence in the time series (i.e., how much the current value of the time series is influenced by its past values), $p$ is the number of lags included in the regression, $\phi_i^{\dagger}$ are coefficients representing the degree of influence of the lagged differences on the current value of the time series, and $\varepsilon_t$ is an error term.

The null hypothesis of the ADF test is that $\rho = 0$, which means that the time series has a unit root and is non-stationary. The alternative hypothesis is that $\rho < 0$, which means that the time series is stationary.

### 3.2.8 Cointegration

The case "where there exist one or more linear combinations of I(1) variables that are I(0) variables" (Nymoen, 2019, p. 340) is called cointegration. One way of testing for cointegration is by the ECM-method. Following the argumentation in (Ericsson & MacKinnon, 2002), the ECM-statistic can be compared to the critical values of the distribution of an ECM-test. The distributions and corresponding tables with rejection values in this paper is adjusting for the model containing a certain number of I(1)-variables. The ECM-statistic is the t-value of the coefficient of the lag of the endogenous variable. The null hypothesis is no relationship between the explanatory and dependent variable, against the alternative of cointegration, i.e., a genuine relationship between random walk variables.

### 3.2.9 Residual misspecification tests

Misspecification testing is an important step when building empirical econometric models. This is because the properties of estimators and of test statistics (i.e., t-values) depend on the assumptions about the distributions of the error-terms. An empirical model is congruent if tests (which use model residuals as data) do not reject the assumption that the model is based on.

Three common misspecification tests are presented in the following. The first test is the AR 1-5 test is a test for the absence of residual autocorrelation using the F-distribution (Harvey, 1981). A rejection of this F-test (i.e., insignificant test results) implies that there is no significant correlation between the error terms in the estimated model equation. The second test is the ARCH (autoregressive conditional heteroskedasticity) 1-4 test, which is a test for a type of heteroskedasticity, i.e., time-dependent variance in the error terms (Engle, 1982). Once more, the F-distribution is used to test the null hypothesis, which, if rejected, points to no model misspecification with regard to residual autoregressive heteroskedasticity. The third and final misspecification test is the Normality test for the normally distributed disturbances (Jarque & Bera, 1980) using the $X^2$-distribution. If the test holds, and thereby the normality assumption, a hypothesis on single parameters in the model can validly be tested using the t-distribution, and the F-distribution can be used to test a joint hypothesis (Nymoen, 2019, p. 91).

## 3.3 Aggregate models and final form equations

The empirical method of this thesis is to construct aggregate macroeconometric models, both univariate and multiple-equation, using the machine learning algorithm Autometrics. The theory and construction of the multiple-equation aggregate model will be the main focus of the thesis. The final form equation (univariate) models will also be made using machine learning, giving another comparable estimation of the economic effect of Covid-19 on GDP.

The definition of a final form equation is that it "expresses a current endogenous variable in terms of exogenous variables and lags of itself, but of no other endogenous variable" (Nymoen, 2019, p. 6). In contrast, a multiple-equation model has more than one endogenous variable in the system. However, there is no logical inconsistency between the two approaches, because each endogenous variable in the multiple-equation model implicitly will have a final form equation (Nymoen, 2023, p. 2).

If a time series is modeled as trend stationary, there will be no long-run effect from Covid-19, according to well-known econometric theory. Likewise, if it is modeled as non-stationary, i.e., as an I(1)-series, the impact of the pandemic will lead to permanent shifts in the level of the variable.

All variables used in my modelling have been modeled as I(1)-series. Through various tests of the null hypothesis of no unit-root(s), such as the ADF-test and Engle-Granger test, it will be shown that the assumption of non-stationary variables is valid. Moreover, the same null hypothesis can normally be rejected for the first-differenced variables. For the final form equations, this implies that any counteracting impacts later in the pandemic period may affect the catch-up in the level of the modelled variable. For the multiple-equation models, this is more of an empirical exercise, as the implications are a lot more complicated (Nymoen, 2023, p. 2).

## 3.4 Autometrics and Impulse Indicator Saturation

Autometrics is a machine learning algorithm that automates variable selection, see Doornik (2009), Hendry & Doornik (2014) and Doornik & Hendry (2018). It seeks to identify an appropriate set of regressors for a given dependent variable. According to Doornik (2009), Autometrics can be viewed as an extension of classical stepwise regression methods that involves a search algorithm for selecting and testing candidate regressors. The algorithm considers a large set of potential regressors which is called the general unrestricted model (GUM). From the GUM, it selects the most relevant variables based on a modified information criterion that accounts for the search process. Its aim is to find the local data generating process (LGDP) using a general-to specific (gets) modelling. One can view the GUM as a search tree containing the whole space of models, where every node contains a unique model, and the GUM being the root of the tree. Moreover, Autometrics can deal with data-dependent modeling issues, such as non-stationarity and endogeneity bias.

Hendry & Doornik (2014) document that Autometrics performs well compared to other regressions methods that involve search algorithms for selecting and testing candidate regressors. There are also other simulation studies by Epprecht et al. (2013) and Muhammadullah et al. (2022) which reach the same conclusion.

Impulse Indicator Saturation (IIS) is a technique used in Autometrics to address the problem of unknown omitted variable bias in a regression model. It extends the GUM that the machine algorithm chooses from by adding a binary indicator variable for each observation. Another interpretation of IIS is that it empirically identifies short-term structural breaks, and that it robustifies the estimated coefficients of the retained economic variables of the models with respect to such breaks (Nymoen, 2019, pp. 400–402). Allowing for IIS to be used may help capture the short-term dynamics of the data without the inclusion of spurious lagged dependent variables (i.e., they are irrelevant explanatory variables in normal times). Moreover, the number of impulse indicators included in the model is determined by the model selection algorithm. The technique has been shown to perform well in simulation studies and in empirical applications (Hendry & Doornik, 2014).

A result of adding impulse indicators for each observation, is that the GUM will consist of more variables than observations. The algorithm offers an elegant solution to this issue, however, using the split-sample IIS algorithm. In the simplest scenario, the indicators are added to the GUM in blocks of T/2, which is feasible since all indicators are uncorrelated. Then, the algorithm adds half of the indicators to the GUM, resulting in the null model in the simplest case. The outcome is recorded, and the first set of indicators is discarded. Next, the second set of T/2 indicators is added, and the selection process is repeated. Finally, the indicators retained by the algorithm are combined and added to the GUM. After completing this process, the selection algorithm can be run again, continuing as if it began with a number of indicators lower than T. For a more detailed explanation, I refer to Nymoen (2019, Chapter 11.7), and Hendry and Doornik (2014, Chapter 15).

## 3.5 Variables

### 3.5.1 Norway

**Y** = GDP Norway, market values, fixed prices, mill. NOK.

**YF** = GDP Mainland-Norway, market values, fixed prices, mill. NOK.

**G** = Public consumption expenditure and gross fixed capital formation in the general government, fixed prices, mill.

**EMI** = Export market indicator, index.

**JOIL** = Gross fixed capital formation (GFCF), oil and gas production and pipeline transportation, and related services, fixed prices, mill. NOK.

**P** = Consumer Price Index (CPI).

**INF** = CPI inflation rate, log approximated ($INF = 100 * p_t - p_{t-4}$, where lowercase p denotes the natural logarithm of CPI).

**RL** = Average interest rate on total bank loans, percent. Equal to $\mathbf{R_L}$ in re-estimation of GDP-equation in NAM-2009.

**RLINF**= Real interest rate, difference between average interest rate on bank loans and log approximated CPI inflation, percent.

**L** = Domestic credit to general public, K2-indicator.

**ARBDAG** = Number of working days per quarter.

**V** = Trade-weighted nominal value of the krone based on import-shares of trading countries.

**P*** = Foreign consumer price index (trade weighted).

**π** = CPI inflation rate ($\pi = \frac{P_t - P_{t-4}}{P_{t-4}} * 100$).

### 3.5.2 Sweden

**Y** = Gross Domestic Product, Total, Constant Prices, Seasonally Adjusted (SA), Market Prices, SEK.

**G** = Central Government Budget, Expenditures, Total, SEK.

**EMI** = Foreign GDP, Closest trading partners (China, Denmark, Euro Area, Norway, Poland, UK, USA), KIX-weighted. Equal to **YF** in re-estimation of GDP-equation in MOSES.

**CPI** = Consumer Price Index (Riksbank Classification), Total, Index.

**RL** = Deposits & Loans, Banks, Lending Rates, By Entity, to Households including NPISH, Period Ending Stock, All Accounts, percent.

**INF** = CPI inflation rate ($INF = 100 * \frac{\Delta_4 CPI_t}{CPI_{t-4}}$). Equal to **π** in re-estimation of GDP-equation in MOSES.

**RLINF** = Real CPI interest rate, percent.

**ARBDAG** = Number of Working Days, Per Quarter.

# Chapter 4: Modelling

## 4.1 Replication of the Norwegian GDP-equation in NAM-2009

To build reliable aggregate models for GDP, I look at what has been previously done in the literature. Over many years, Bårdsen and Nymoen has been operating and developing an empirical dynamic macroeconometric model of the Norwegian economy. The project is called Norwegian Aggregate Model (NAM) (https://normetrics.no/nam/, 2022). The current model version has become quite large in terms of equations and variables. Hence, the operational version of NAM represents way more relationships of the Norwegian economy than earlier. Also, for a focused research question like mine, i.e., the magnitude and significance of covid effects on GDP and stability of "main" relationships in the first period after the coronavirus crisis, the full model is cumbersome to use. In addition, it requires specialized software to operate, which is unavailable to me. Moreover, there exists no comparable model for Sweden, and since the aim is to analyze both Norway and Sweden, a feasible approach is to create small, aggregate models for both countries. Hence, for the purposes of this thesis, I move away from using the operational NAM as anything more than a reference point for my own modelling.

However, as a starting point, I took a closer look at an older version dating back to 2009. This was an aggregate model in the true letter of the word, see Bårdsen & Nymoen (2009). As a first step, I re-estimated the GDP-equation in NAM-2009 using the current vintage of time series data. Because of data revisions, the aim is not exact replication, but to establish a reference point for my own modelling. The credit indicator that I need in my own modelling work, L, is only available as a time series dating back to the second quarter of 1988, which reduced the sample size by eight quarters compared to NAM-2009.

As a reference, I include the GDP-equation in NAM-2009 (equation number (17.44) in Bårdsen & Nymoen (2009)):

$$\Delta y_t = -\underset{(0.041)}{0.21}\left[y_{t-2} - 0.9g_{t-1} - 0.16(v + p^* - p)_{t-1} + 0.006(R_L - \pi)_{t-1} - \mu_y\right] - \underset{(0.091)}{0.74}\Delta y_{t-1} + \underset{(0.058)}{0.42}\Delta g_t + \underset{(0.11)}{0.67}\Delta(l-p)_{t-1} \tag{4.1}$$

In equation (4.1) above, $\mu_y$ denotes the mean of the long-run relationship, $g$ is public consumption expenditure, and $y$ is total GDP in Norway. The expression $v + p^* - p$ denotes the real exchange rate, $R_L - \pi$ denotes the real interest rate, and $l - p$ is an indicator for real credit. Moreover, in this equation, as well as all following equation through the thesis, estimated standard errors are given in round brackets below the coefficients, lowercase of a variable denotes the natural logarithm of that variable (e.g., y = ln(Y)), and Δ denotes the first difference.

For comparison, below follows the re-estimation of the GDP-equation on ECM-form using updated data:

$$\Delta y_t = -\underset{(0.07)}{0.48}[\mathrm{y}_{t-2} - 0.9\mathrm{g}_{t-1} - 0.16(\mathrm{v}+\mathrm{p}^* - \mathrm{p})_{t-1} + 0.006(\mathrm{R_L} - \pi)_{t-1}] \\ -\underset{(0.09)}{0.48}\Delta y_{t-1} + \underset{(0.04)}{0.72}\Delta g_t + \underset{(0.18)}{0.73}\Delta(l-p)_{t-1} + \underset{(0.20)}{1.23} \tag{4.2}$$

| | | |
|---|---|---|
| OLS | Sample: 1998(2) - 2007(1) | N = 76 |
| $AR_{1-5}$: | F(5,66) = 2.1778 [0.0671] | $\hat{\sigma}100 = 0.18$ |
| $ARCH_{1-4}$: | F(4,68) = 0.6872 [0.6032] | RSS = 0.0025 |
| Normality: | $\chi^2(2)$ = 1.5089 [0.4703] | |

The re-estimation of the GDP-equation on unrestricted form gave the following result:

$$\Delta y_t = -\underset{(0.08)}{0.55}y_{t-2} + \underset{(0.08)}{0.48}g_{t-1} + \underset{(0.04)}{0.02}(v+p^*-p)_{t-1} - \underset{(0.0015)}{0.004}(R_L - \pi)_{t-1} \\ -\underset{(0.10)}{0.55}\Delta y_{t-1} + \underset{(0.04)}{0.71}\Delta g_t + \underset{(0.26)}{1.05}\Delta(l-p)_{t-1} + \underset{(0.49)}{1.58} \tag{4.3}$$

| | | |
|---|---|---|
| OLS | Sample: 1998(2) - 2007(1) | N = 76 |
| $AR_{1-5}$: | F(5,63) = 2.2135 [0.0638] | $\hat{\sigma}100 = 0.18$ |
| $ARCH_{1-4}$: | F(4,68) = 0.8788 [0.4813] | RSS = 0.0023 |
| Normality: | $\chi^2(2)$ = 2.5087 [0.2853] | |

Below equation (4.2) and (4.3) above, the results from the misspecification tests presented in chapter 3.2.9 are given, as well as the residual percentage standard deviation ($\hat{\sigma}100$) and residual sum of squares (RSS). The residual percentage standard deviation can be directly interpreted as errors of the dependent variable's residuals. The RSS on the other hand, measures the level of variance in the error terms, meaning that as RSS becomes smaller, the tighter will the fit of the model be to the data. The misspecification test results are all rejecting the null hypothesis of potential model misspecification, and the low estimated $\hat{\sigma}100$ and RSS implies a good fit of the data, with small estimated standard errors.

To check whether the ECM-term in the re-estimated GDP-equation from NAM-2009 contains cointegrating parameters, I perform an ECM-test for no cointegration. The t-value of relevance is the reported t-value for the lagged dependent variable, $y_{t-2}$, in the unrestricted estimation. The t-value of -6.73 is rejected against the critical value at 1% of -4.09 from table 3 in Ericsson & MacKinnon (2002), due to three I(1)-variables, corresponding to N=3 in their notation.

**Table 4.1:** Values of the coefficients in equations (4.1)-(4.3)

| Variable | NAM-2009 | Re-estimated, ECM-form | Re-estimated, unrestricted |
|---|---|---|---|
| $y_{t-2}$ | -0.21 | -0.48 | -0.55 |
| $g_{t-1}$ | -0.9 | -0.9 | 0.48 |
| $(v + p^* - p)_{t-1}$ | -0.16 | -0.16 | 0.02 |
| $(RL - \pi)_{t-1}$ | 0.006[1] | 0.006 | -0.004 |
| $\Delta y_{t-1}$ | -0.74 | -0.48 | -0.55 |
| $\Delta g_t$ | 0.42 | 0.72 | 0.71 |
| $\Delta(l - p)_{t-1}$ | 0.67 | 1.23 | 1.58 |

The GDP-equation in NAM is first estimated using OLS, then re-estimated using ECM. In both cases the magnitudes of the parameters are larger than in the original equation, but they are similar, and the coefficient signs are correct. This coincided with the beforehand expectation.

[1] Corrected typo in equation (17.44) in Bårdsen & Nymoen (2009) which reported the coefficient as 0.06.

## 4.2 Replication of the Swedish GDP-equation in MOSES

For the Swedish macroeconomy, I base my modelling on MOSES, an aggregate macroeconometric model of the Swedish economy which dates back to 2011. This was a similar project as NAM, developed by Bårdsen, den Reijer, Jonasson and Nymoen, in cooperation with the Swedish Riksbank. The size of MOSES, in terms of included variables and number of difference equations, is comparable to NAM-2009. This makes it suitable as a starting point of my own aggregate model of the Swedish economy. In the same way as for the Norwegian data, I first re-estimated the GDP-equation with the same specification as in the MOSES documentation. This gives a reference point for my own model. However, for the same reason as with NAM-2009, I cannot expect anything close to an exact replication, and differences must be expected due to data revisions.

Another issue is that data on the original YF-variable is not available, so it will have to be reconstructed and likely differ from its original. YF is a variable for trade weighted foreign GDP, which is interpretable as an indicator of the growth in Swedish export markets. I have constructed the variable by the use of growth in countries that are trade partners of Sweden. The GDP-values are weighted by a trade indicator the Riksbanken (The Swedish Central Bank) produces. In the original MOSES GDP-equation, the index used was TCW, or "Total Competitiveness Weights." However, the TCW index was discontinued in 2021, and it was replaced by an improved version called the "Krona index", or just KIX. It is an index where "the weights are based on total flows of processed goods and commodities for 32 countries" (Riksbanken, 2023), while taking imports and exports into account.

I constructed the variable using the following method: First since, the KIX is an annual index, the weights were divided by four to match the quarterly GDP data. Second, I simplified somewhat by choosing only the most important trade partners, the cutoff being an average KIX-weight above 4 percentage points. This resulted in seven countries, including the Euro area. I then took the logarithmic differences of the collected quarterly GDP data and multiplied these with the respective KIX-weights. For each quarter, the KIX-weighted GDP growth was summed. Finally, choosing one as the initial value, the summed trade-weighted GDP growth data is added cumulatively, resulting in the YF-variable used in the specification of the new empirical GDP model equation.

Finally, another issue is represented by the structural shift dummies denoted $x_{st}$ (proxying a shift in export shares from 2004(1)), which will not be included in the re-estimation because they are not needed in the data series I have used. Nevertheless, the re-estimated MOSES GDP-equation is interesting as a reference point for my own work with specifying a GDP-equation. In the equations below, *Y* is GDP, *YF* is the foreign market indicator, *G* denotes government expenditure and $RL - \pi$ denotes the real interest rate. As before, lowercase letters is used to signify the natural logarithm of a variable, and Δ the first difference.

GDP-equation in the MOSES documentation (eq. (53) in (Bårdsen et al., 2011):

$$\Delta y_t = -\underset{(0.055)}{0.18}\,(y-yf)_{t-1} + \Delta\left[\left(\frac{G}{Y}\right)_{t-1} g_t\right] - \underset{(0.13)}{0.38}\Delta y_{t-1} - \underset{(0.0008)}{0.0014}\,(RL-\pi)_{t-4} + \underset{(0.41)}{3.4}\,x_{s_t \Delta yf_t} + \underset{(0.4)}{1.4}\,x_{s_{t-2}\Delta yf_{t-2}} + \underset{(0.49)}{1.5} \tag{4.4}$$

Re-estimated GDP-equation using same sample size as in MOSES, 1997(1) – 2009(4):

$$\Delta y_t = -\underset{(0.011)}{0.01}\,(y-yf)_{t-1} - \underset{(0.0009)}{0.001}\,\Delta\left[\left(\frac{G}{Y}\right)_{t-1} g_t\right] + \underset{(0.14)}{0.04}\Delta y_{t-1} + \underset{(0.0015)}{0.001}\,(RL-\pi)_{t-4} + \underset{(0.05)}{0.12}\Delta yf_t + \underset{(0.04)}{0.05}\Delta yf_{t-2} + \underset{(0.29)}{0.34} \tag{4.5}$$

| | | |
|---|---|---|
| OLS | Sample: 1997(1) - 2009(4) | N = 52 |
| $AR_{1-5}$: | F(4,41) = 1.0719 [0.3828] | $\hat{\sigma}100 = 0.85$ |
| $ARCH_{1-4}$: | F(4,44) = 0.1517 [0.9612] | RSS = 0.0032 |
| Normality: | $\chi^2(2)$ = 16.660 [0.0002]** | |

Re-estimated GDP-equation using a longer sample, which is possible with the data set that I have collected for the thesis, 1996(1) – 2020(4):

$$\Delta y_t = -\underset{(0.008)}{0.03}(y - yf)_{t-1} - \underset{(0.001)}{0.0015}\,\Delta\left[\left(\frac{G}{Y}\right)_{t-1} g_t\right] - \underset{(0.08)}{0.25}\Delta y_{t-1} + \underset{(0.001)}{0.003}(RL - \pi)_{t-4} + \underset{(0.03)}{0.23}\Delta yf_t + \underset{(0.039)}{0.03}\,\Delta yf_{t-2} + \underset{(0.22)}{0.92} \qquad (4.6)$$

| | | |
|---|---|---|
| OLS | Sample: 1996(1) - 2020(4) | N = 100 |
| $AR_{1-5}$: | F(5,88) = 2.9099 [0.0177]* | $\hat{\sigma}100 = 1.18$ |
| $ARCH_{1-4}$: | F(4,92) = 0.8170 [0.5175] | RSS = 0.0131 |
| Normality: | $\chi^2(2)$ = 30.997 [0.0000]** | |

As expected, there are not only trivial differences when the GDP-equation in MOSES is re-estimated. When using the same sample size as in the original equation, the parameters' magnitudes are generally smaller, see equation (4.5). When the sample is increasing with another 12 years as in equation (4.6), the re-estimation improves considerably. The magnitudes of the parameters increase towards those in equation (4.4), and the coefficient signs in the main the same is in the MOSES documentation.

Moreover, the misspecification tests reported below equations (4.5) and (4.6) show signs of misspecification in both cases. For both equations, the normality test is highly significant, while the test for autoregression becomes significant at the 5% significance level in equation (4.6). Because the equation above only are re-estimations, the misspecifications will not be considered further, but note that my own estimations hopefully will perform better.

While the estimated coefficients in the re-estimated GDP-equations in NAM-2009 and MOSES are different from the original estimations due to updated data and other reasons mentioned above, they are still similar and relevant for further estimation. Common for them both is that they are estimated using explanatory variables from the aggregate demand side of the economy. The GDP-equations will in the following be used as reference points when creating my own framework for aggregate models of the Norwegian and Swedish economy.

## 4.3 Aggregate model of the Norwegian economy

To create a new model of the Norwegian economy, I began by estimating a dynamic GDP-equation for Mainland-Norway. The term for Mainland-Norway GDP is adopted from Statistics Norway (2023): It refers to GDP without value added in oil and gas extraction, pipeline transportation and ocean transport. In the work with the specification of the model equation, I found the machine learning algorithm Autometrics with impulse indicator saturation (IIS) useful. The algorithm and IIS were explained in chapter 3.4, where I also showed that the algorithm is well-documented and has shown to perform well.

### 4.3.1 Specifying an equation for Norwegian Mainland GDP using Autometrics

I began the modelling of a Norwegian GDP-equation by using the machine learning algorithm Autometrics to choose significant explanatory variables. The general unrestricted model (GUM) consisted of carefully selected variables that typically are relevant when modelling GDP. The re-estimation of the GDP-equation in NAM-2009 was also a source of inspiration when choosing potentially relevant variables. Included in the GUM, were the following variables, with lag lengths varying from 1 to 4:

$\Delta yf$ and $yf$, which is crucial to include as current GDP is likely to be correlated with lags of GDP. The export market indicator, $emi$, was included because GDP in Norway typically is correlated with general market growth. Net government expenditure, $g$ and $joil$, total expenditure in the oil sector, are included as fluctuations in these may be important explanatory factors for GDP. The real interest rate, $RLINF$, was included as it is an important explanatory variable for economic activity. Furthermore, the credit indicator that was found highly significant in NAM-2009, $\Delta(l-p)_{t-1}$ was included in the GUM. Additionally, an indicator for the number of working days per quarter, $arbdag$, as well as seasonal dummies denoted $CS_t$ were included to cover seasonal variations. Finally, by using the IIS method in Autometrics explained above, impulse indicator variables for each observation were added to the GUM.

One potential caveat in automatic variable selection algorithms is that repeated testing can lead to inflated Type-I error probability levels (Nymoen, 2023). In Autometrics, the “Target size” is the user’s choice of overall significance level. Using Monte Carlo experiments, Castle et al.

(2012) showed that the probability of retaining any irrelevant dummies are close to the target size ($\alpha$) set. However, the probability of Type-II error will necessarily increase, as setting a tight target size may reduce the probability of detecting true break periods in the sample.

Choosing the target size of 1 percent, Autometrics reported a final model where $yf_{t-1}$ was retained with a t-value of -6.44. The final model equation for Mainland-GDP estimated using Autometrics was:

$$\begin{aligned}\Delta yf_t = &- \underset{(0.04)}{0.36}\, \Delta yf_{t-1} + \underset{(0.03)}{0.09}\, \Delta yf_{t-4} - \underset{(0.02)}{0.17} yf_{t-1} + \underset{(0.02)}{0.08}\, g_{t-1} + \underset{(0.01)}{0.05}\, emi_{t-1} \\ &+ \underset{(0.03)}{0.41}\, \Delta arbdag_t + \underset{(0.05)}{0.26}\, arbdag_{t-1} + + \underset{(0.03)}{0.15}\, \Delta g_t + \underset{(0.03)}{0.09}\, \Delta g_{t-1} \\ &+ \underset{(0.005)}{0.015}\, \Delta joil_t + \underset{(0.005)}{0.022}\, \Delta joil_{t-1} + \underset{(0.005)}{0.021}\, \Delta joil_{t-2} + \underset{(0.05)}{0.22}\, \Delta emi_t \\ &- \underset{(0.004)}{0.014}\, CS_t + \underset{(0.004)}{0.028}\, CS_{t-2} - \underset{(0.009)}{0.028}\, I_{1991(4)} + \underset{(0.009)}{0.034}\, I_{2006(4)} \end{aligned} \tag{4.7}$$

| | | |
|---|---|---|
| OLS | Sample: 1988(2) – 2019(4) | N = 127 |
| AR 1-5 test: | F(5,105) = 1.4208 [0.2228] | $\hat{\sigma}100 = 0.88$ |
| ARCH 1-4 test: | F(4,119) = 1.4169 [0.2325] | RSS = 0.0086 |
| Normality test: | $\chi^2(2)$ = 1.1825 [0.5536] | |

Just like previously, estimated standard errors are reported in brackets below the estimated coefficients in equation (4.7). $CS_t$ are denoting centered quarterly seasonal dummies in period t, and $I_{1991(4)}$ is annotation for an indicator variable for the last quarter of 2006. It is worth noting that neither $RLINF$ or $\Delta(l-p)_{t-1}$ was retained as explanatory variables after Autometrics' selection process. The historic variation that would have been explained by either of the two mentioned variables is explained by more significant retained variables. Furthermore, Autometrics chose to include two impulse indicators: $I_{1991(4)}$, which is in a period of a well-known Norwegian banking crisis, and $I_{2006(4)}$, a quarter with unusually high economic activity. The final model equation did not show any signs of misspecification in the tests, and the residual percentage standard deviation as well as the RSS indicate a tight model fit to the data, with small estimated error terms.

The t-value of the $yf_{t-1}$ is the so-called ECM-statistic of the ECM-test of no cointegration. The 1 percent critical value of the ECM-test of Ericsson and MacKinnon (2002) with three I(1) variables (corresponding to N=3 in table 3 in their notation) is -3.79. Hence, the t-value of -6.44 rejects the null of no cointegration. Based on this result I constructed the following ECM-variable:

$$yf_{ECM_t} = yf_t - \frac{\hat{\gamma}_1}{\hat{\pi}} g_t - \frac{\hat{\gamma}_2}{\hat{\pi}} emi_t \qquad (4.8)$$

The "hat-notation" in (4.8) is used to represent the corresponding estimated coefficient in (4.7), e.g., $\gamma_1$ represents the coefficient for $g_t$. Inserting for the estimated coefficients yielded the following ECM-equation:

$$yf_{ECM_t} = yf_t - 0.47\ g_t - 0.31\ emi_t \qquad (4.9)$$

Furthermore, the plot of $yf_{ECM}$ is added below. The plot shows graphically that a plausible interpretation of the variable is that it is in fact a stationary variable, i.e., an I(0) series.

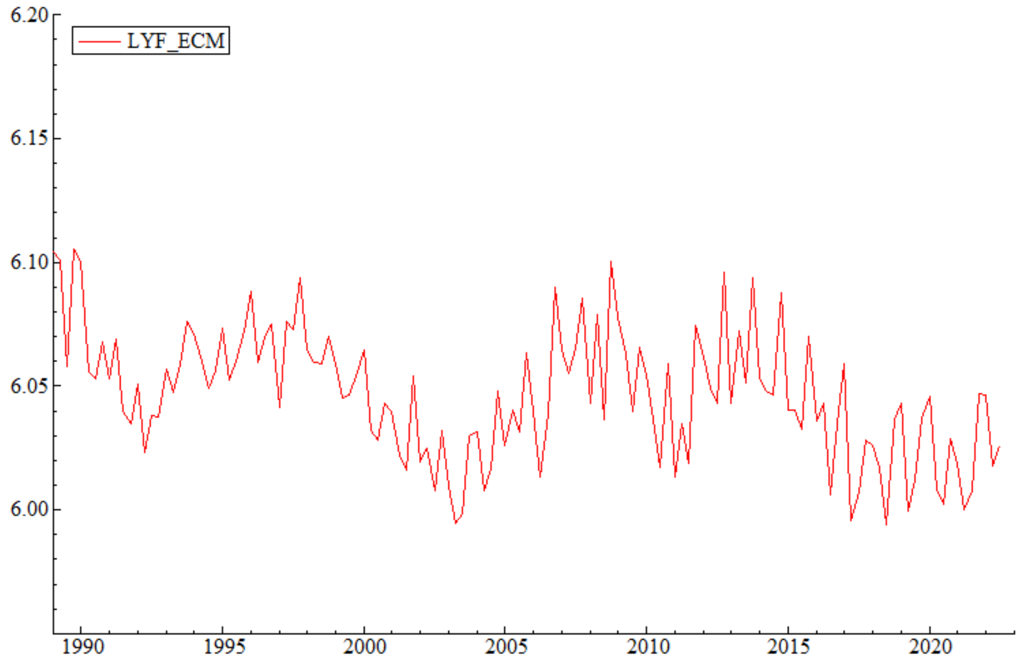


**Figure 4.1:** Plot of the $yf_{ECM}$-variable defined by equation (4.9)

Sample: 1988(2) – 2022(3)

Before continuing, I performed a robustness check of cointegration. The variables used to construct the ECM-equation, namely $yf_{t-1}$, $g_{t-1}$ and $emi_{(t-1)}$, were retained in equation (4.7). G and EMI are two important demand indicators for economic activity, affecting YF both directly and indirectly (through other channels). Using a static regression to estimate a potential significant long-run relationship between these variables, I tested whether these variables were cointegrated using the Engle-Granger method. The estimated equation became:

$$yf_t = 0.45\ g_t + 0.30\ emi_t + 6.33 \tag{4.10}$$

In order to test the interpretation that (4.10) is a valid long-run relationship, I took the estimated coefficients as given and define the long-run residuals as the variable $yf_{static}$:

$$yf_{static_t} = yf_t - 0.45 g_t - 0.30\ emi_t \tag{4.11}$$

The estimated coefficients in the long-run static regression are almost identical to the estimated residuals in the ECM-equation (4.9), implying a close correspondence, and followingly, a robust cointegrated relationship.

To investigate the hypothesis that also the long-run static regression is cointegrated, I tested the null hypothesis of a unit-root using an Augmented Dickey-Fuller (ADF) test. The ADF-test is another method of testing the null of no cointegration, like the ECM-test previously performed. The ADF-test takes into account that the series is a residual for estimation. Therefore, the critical values that this cointegration test uses depend on the number of variables in the "cointegration regression," which is three in this case. The critical value can be found in table 1 in MacKinnon (1990), where the critical value for N=3 and no trend variable at 5% is -3.74. This means that, based on the values in the t-adf column below, the null hypothesis of a unit-root (i.e., no cointegration) can be rejected at the 1% critical level.

In table 4.2 below, the critical values are from MacKinnon (1990), which improves on previous table with critical values for the Engle-Granger cointegration test. Two stars next to the values in the t-adf-column means that it is significant at the 1% significance level, and one star at the 5% significance level. Moreover, it can be noted that up to the ninth differenced lag of the $yf_{static}$ variable, the p-value (denoted t-prob in the table) is significant. This implies that the ninth lag is the smallest model that would have given valid test results.

**Table 4.2:** ADF-test of $yf_{static}$

| | | | |
|---|---|---|---|
| Critical values from MacKinnon (1990): 5% = -3.74, 1% = -4.29 (N=127, Constant) | | | |
| Sample: 1988(2) – 2019(4) | | | |
| D-lag | t-adf | t-DY_lag | t-prob |
| 10 | -4.981** | 1.776 | 0.0784 |
| 9 | -4.680** | 2.034 | 0.0442 |
| 8 | -4.406** | 2.766 | 0.0066 |
| 7 | -4.202* | -2.448 | 0.0158 |
| 6 | -4.308** | 2.599 | 0.0105 |
| 5 | -4.064* | 3.707 | 0.0003 |
| 4 | -4.114* | 2.575 | 0.0112 |
| 3 | -4.285* | -6.393 | 0.0000 |
| 2 | -4.570** | -2.386 | 0.0186 |
| 1 | -5.324** | -3.639 | 0.0004 |
| 0 | -7.089** | | |

Through various cointegration tests, I now have confirmed that I have a model consisting of only stationary variables and chose to re-run Autometrics with $yf_{ECM_{t-1}}$ in the GUM instead of the three lagged level terms. The resulting model was:

$$\begin{aligned}\Delta yf_t = &- \underset{(0.02)}{0.40}\,\Delta yf_{t-1} + \underset{(0.02)}{0.17}\,\Delta yf_{t-4} - \underset{(0.02)}{0.20}\,yf_{ECM_{t-1}} + \underset{(0.02)}{0.45}\,\Delta arbdag_t + \underset{(0.03)}{0.30}\,arbdag_{t-1} \\ &+ \underset{(0.02)}{0.09}\,\Delta g_t + \underset{(0.005)}{0.023}\,\Delta joil_t + \underset{(0.005)}{0.028}\,\Delta joil_{t-1} + \underset{(0.005)}{0.020}\,\Delta joil_{t-2} + \underset{(0.04)}{0.23}\,\Delta emi_t \\ &+ \underset{(0.003)}{0.037}\,CS_{t-2} + \underset{(0.009)}{0.032}\,I_{1996(4)} - \underset{(0.008)}{0.019}\,I_{1998(2)} - \underset{(0.008)}{0.023}\,I_{2001(3)} \\ &+ \underset{(0.008)}{0.032}\,I_{2006(4)} + \underset{(0.008)}{0.022}\,I_{2007(4)} - \underset{(0.008)}{0.025}\,I_{2011(1)} + \underset{(0.008)}{0.017}\,I_{2012(2)} \end{aligned} \tag{4.12}$$

| | | |
|---|---|---|
| OLS | Sample: 1988(2) - 2019(4) | N = 127 |
| $AR_{1-5}$: | F(5, 104) = 1.3951 [0.2322] | $\hat{\sigma}100 = 0.83$ |
| $ARCH_{1-4}$: | F(4, 119) = 0.4679 [0.7591] | RSS = 0.0076 |
| Normality: | $\chi^2(2)$ = 1.0281 [0.5981] | |

More impulse indicators have now been included, while $\Delta g_{t-1}$ and $CS_t$ has been excluded in the re-estimation. Other than that, the variables in the final model equation with the ECM-variable (4.12) are equal to those in equation (4.7). The impulse indicator variables explain variation in the historic data which is not explained by the retained explanatory variables in the model equation. The inclusion of more impulse indicators is not an issue regarding estimation consistency, as explained in the paragraph about Impulse Indicator Saturation.

Both the residual percentage standard deviation ($\hat{\sigma}100$) and the residual sum of squares (RSS) has decreased slightly compared to before, which indicate an improved fit of the model to the data. This follows from the general rule that as more significant explanatory variables are added to the model, the overall fit of the model will be improved. This is because the added variables are able to explain some of the variation in the dependent variable that was previously unexplained. As an example, the variable for government expenditure, G, is an important explanatory variable for the economic activity, such that adding this variable to the equation will reduce both the residual percentage standard deviation and the RSS. Because I have used Autometrics to retain only significant explanatory variables, these will all be contributing to improving the fit of the data. Additionally, the misspecification tests $AR_{1-5}$, $ARCH_{1-4}$ and $Normality$ are reported with their respective p-values, all insignificant.

The plot of the $\Delta yf$-variable is included below, see figure 4.2. It shows the good fit of the estimation of Mainland-Norway GDP graphically. Moreover, recursive plots of the coefficient (with standard errors) on the lagged ECM-variable ($yf_{ECM_{t-1}}$), as well as 1-step residuals, 1-step Chow tests and break-point Chow tests for this variable are included (see figure 4.3).

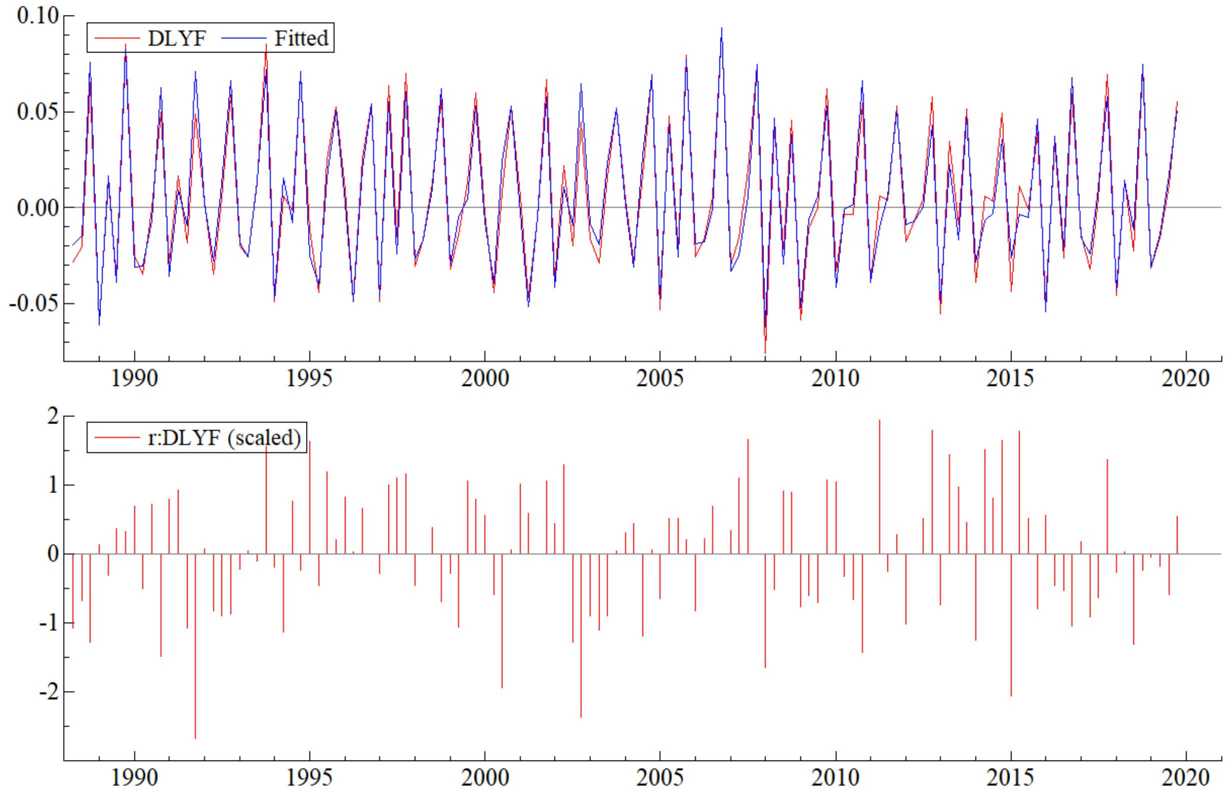


**Figure 4.2:** Actual and fitted values of $\Delta yf$ (DLYF in figure) in equation (4.12), and scaled residuals (DLYF scaled) in lower panel

Recursive estimation is a practical way to investigate the stability of the parameters in a model equation (Nymoen, 2019, p. 95). This can be performed when the data are time series because there is a natural ordering of the observations. In practice, the model equation is first fitted to an initial sample $t = 1,2, \dots, T_1$, and then subsequently fitted to samples of $T_1 + 1, t_1 + 2$ up to $T$ observations. The graphs of recursive estimations are a powerful way to study parameter constancy and stability because the absence of constancy will be easily recognized.

The recursive test of the coefficient on $yf_{ECM_{t-1}}$ represents the estimated effect of this variable on the dependent variable. That the confidence intervals of this coefficient are moving quickly away from zero implies that it has a statistically significant effect on the dependent variable at the 5% significance level. Moreover, the residual 1-step test recursively estimates the model and test the residuals' statistical significance. The plot of one period ahead ("1-up") Chow tests and the break-point ("Ndn") Chow tests can be inspected for signs of significant changes (structural breaks in the relationship model equation. These tests are included to graphically show the parameter constancy and stability of the ECM-variable, which in this case, all points to good parameter constancy.

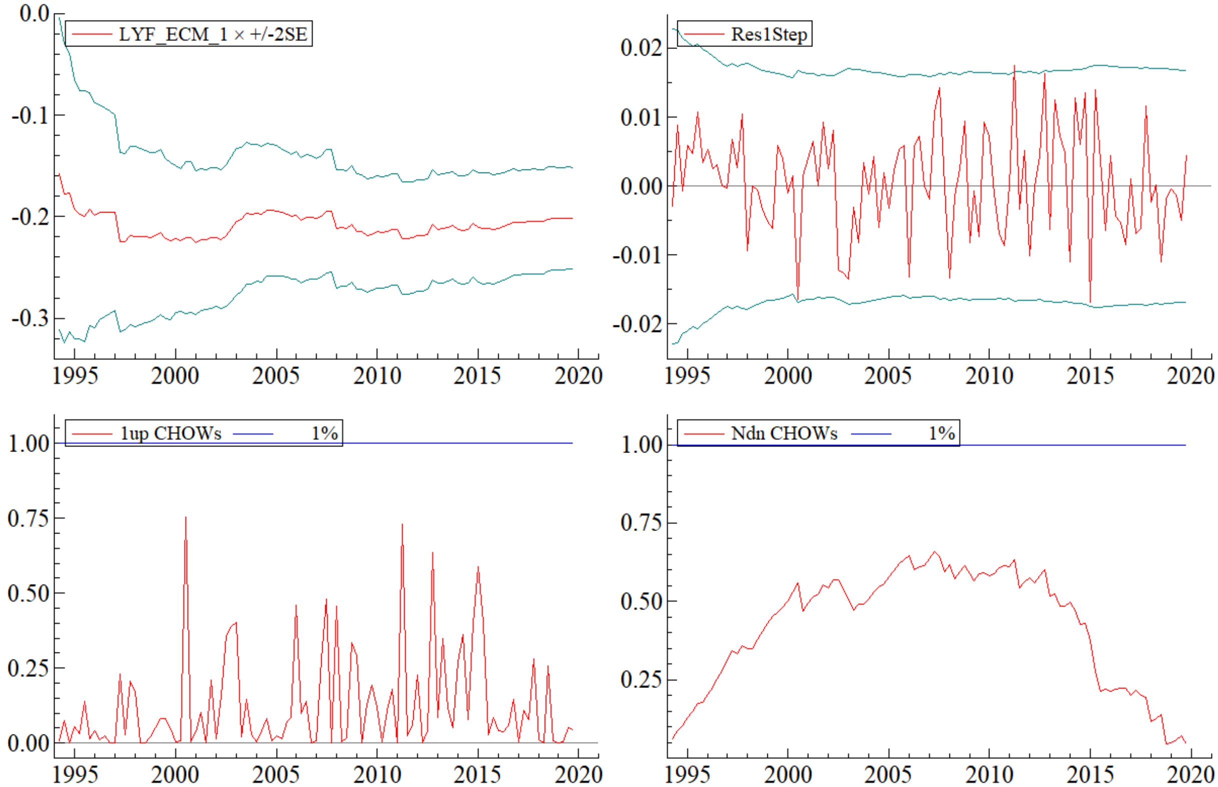


**Figure 4.3:** Recursive parameter constancy tests of $yf_{ECM_{t-1}}$ in equation (4.12)

### 4.3.2 The estimated marginal equations in the Norwegian equation system

Continuing with GDP-equation (4.12) as the conditional model in the equation system, I move on to estimating the marginal models. For the purpose of this thesis, which is to create a small, aggregate model of the Norwegian economy, some simplifications are done in this process. The marginal models will be constructed using only lags of itself, deterministic variables, and impulse indicators, resulting in autoregressive processes. In equation (4.12), there are three endogenous variables included as explanatory variables, namely $\Delta g$, $\Delta emi$ and $\Delta joil$, which will need marginal models to complete an equation system.

I began by modelling $\Delta g$, including six lags of the dependent variable, the deterministic variable *arbdag* (in trend form and differenced form) and centered quarterly seasonal dummies in the GUM. In addition, by allowing for IIS when using Autometrics, impulse indicators for each period in the sample is also included. Then I repeated this process for $\Delta emi$ and $\Delta joil$, replacing the lags of the dependent variable with the respective variable which I was estimating each time. Moreover, because all these three variables are economic indicators with a long-run trend, constant terms were also fixed in the estimations. The estimated coefficients of the retained variables in the marginal models are summarized in the table below. Also included are the misspecification tests for the marginal models, with no indication of any model misspecification.

**Table 4.3:** Estimated coefficients in the marginal models of Norwegian model

| $\Delta g$ | | $\Delta emi$ | | $\Delta joil$ | |
|---|---|---|---|---|---|
| Variable name | Coefficient | Variable name | Coefficient | Variable name | Coefficient |
| $Constant$ | 0.019 (0.002) | $Constant$ | 0.009 (0.001) | $Constant$ | 0.015 (0.007) |
| $\Delta g_{t-1}$ | −0.95 (0.05) | $\Delta emi_{t-1}$ | 0.34 (0.06) | $\Delta joil_{t-1}$ | −0.37 (0.04) |
| $\Delta \mathrm{g}_{\mathrm{t-2}}$ | −0.68 (0.05) | $I_{2008(4)}$ | −0.06 (0.01) | $\Delta joil_{t-4}$ | 0.15 (0.04) |
| $\Delta g_{t-3}$ | −0.56 (0.05) | $I_{2009(1)}$ | −0.07 (0.01) | $CS_t$ | −0.14 (0.08) |
| $\Delta arbdag$ | 0.15 (0.02) | $I_{2009(3)}$ | 0.03 (0.01) | $I_{1988(4)}$ | 0.22 0.08 |
| $I_{1991(4)}$ | 0.06 (0.01) | | | $I_{1989(1)}$ | −0.25 (0.08) |
| $I_{2007(4)}$ | 0.04 (0.01) | | | $I_{1990(2)}$ | −0.35 (0.08) |
| $I_{2008(3)}$ | 0.06 (0.01) | | | $I_{1991(3)}$ | 0.22 (0.08) |
| $I_{2009(4)}$ | 0.03 (0.01) | | | $I_{1994(4)}$ | −0.29 (0.08) |
| $I_{2010(4)}$ | −0.04 (0.01) | | | $I_{1996(1)}$ | 0.31 (0.08) |
| $I_{2011(4)}$ | −0.04 (0.01) | | | $I_{1996(3)}$ | −0.50 (0.08) |
| $I_{2012(1)}$ | −0.03 (0.01) | | | $I_{1996(4)}$ | −0.55 (0.08) |
| | | | | $I_{1997(1)}$ | 0.44 (0.08) |
| | | | | $I_{1997(2)}$ | 0.40 (0.08) |
| | | | | $I_{1999(4)}$ | −0.54 (0.08) |
| | | | | $I_{2000(1)}$ | 0.37 (0.08) |
| AR 1-5 test: | | AR 1-5 test: | | AR 1-5 test: | |
| F(5,110) = 0.3132 [0.9041] | | F(5,117) = 3.121 [0.0111]* | | F(5,106) = 0.8513 [0.5165] | |
| ARCH 1-4 test: | | ARCH 1-4 test: | | ARCH 1-4 test: | |
| F(4,119) = 0.4295 [0.7871] | | F(4,119) = 0.4777 [0.7520] | | F(4,119) = 1.6568 [0.1646] | |
| Normality test: | | Normality test: | | Normality test: | |
| $\mathrm{X}^2(2)$ = 0.9459 [0.6231] | | $\mathrm{X}^2(2)$ = 1.1352 [0.5669] | | $\mathrm{X}^2(2)$ = 0.4205 [0.8104] | |

OLS, Sample: 1988(2) – 2019(4), N=127

### 4.3.3 The multiple-equation model for Mainland-Norway GDP

The individual equations were collected to form an equation system by estimating a SEM (simultaneous equation model) with the equation for $\Delta yf$ as the conditional model and the equations for $\Delta g$, $\Delta emi$ and $\Delta joil$ as marginal models. The resulting equation system is presented below, where (4.13) is the conditional model and (4.14) - (4.16) are the marginal models:

$$\begin{aligned}\Delta yf_t = &- \underset{(0.02)}{0.41}\,\Delta yf_{t-1} + \underset{(0.02)}{0.18}\,\Delta yf_{t-4} - \underset{(0.02)}{0.19}\,yf_{ECM_{t-1}} + \underset{(0.02)}{0.45}\,\Delta arbdag_t \\ &+ \underset{(0.03)}{0.28}\,arbdag_{t-1} + \underset{(0.03)}{0.06}\,\Delta g_t + \underset{(0.006)}{0.021}\,\Delta joil_t + \underset{(0.005)}{0.026}\,\Delta joil_{t-1} + \underset{(0.005)}{0.020}\,\Delta joil_{t-2} \\ &+ \underset{(0.06)}{0.25}\,\Delta emi_t + \underset{(0.003)}{0.036}\,CS_{t-2} + \underset{(0.009)}{0.030}\,I_{1996(4)} - \underset{(0.008)}{0.019}\,I_{1998(2)} - \underset{(0.008)}{0.024}\,I_{2001(3)} \\ &+ \underset{(0.008)}{0.034}\,I_{2006(4)} + \underset{(0.008)}{0.023}\,I_{2007(4)} - \underset{(0.008)}{0.025}\,I_{2011(1)} + \underset{(0.008)}{0.017}\,I_{2012(2)} \end{aligned} \quad (4.13)$$

$$\begin{aligned}\Delta g_t = &\underset{(0.002)}{0.019}\,Constant - \underset{(0.04)}{0.85}\Delta g_{t-1} - \underset{(0.05)}{0.68}\Delta g_{t-2} - \underset{(0.05)}{0.55}\Delta g_{t-3} + \underset{(0.02)}{0.15}\Delta arbdag \\ &+ \underset{(0.01)}{0.07}\,I_{1991(4)} + \underset{(0.01)}{0.04}\,I_{2007(4)} + \underset{(0.01)}{0.06}\,I_{2008(3)} + \underset{(0.01)}{0.03}\,I_{2009(4)} \\ &- \underset{(0.01)}{0.04}\,I_{2010(4)} - \underset{(0.01)}{0.04}\,I_{2011(4)} - \underset{(0.01)}{0.03}\,I_{2012(1)} \end{aligned} \quad (4.14)$$

$$\begin{aligned}\Delta emi_t = &\underset{(0.001)}{0.008}\,Constant + \underset{(0.06)}{0.34}\,\Delta emi_{t-1} - \underset{(0.01)}{0.06}\,I_{2008(4)} - \underset{(0.01)}{0.07}\,I_{2009(1)} \\ &+ \underset{(0.01)}{0.03}\,I_{2009(3)} \end{aligned} \quad (4.15)$$

$$\begin{aligned}\Delta joil_t = &\underset{(0.007)}{0.015}\,Constant - \underset{(0.04)}{0.38}\,\Delta joil_{t-1} + \underset{(0.04)}{0.15}\,\Delta joil_{t-4} - \underset{(0.08)}{0.14}\,CS_t \\ &+ \underset{(0.07)}{0.23}\,I_{1988(4)} - \underset{(0.08)}{0.24}\,I_{1989(1)} - \underset{(0.07)}{0.35}\,I_{1990(2)} + \underset{(0.07)}{0.21}\,I_{1991(3)} \\ &- \underset{(0.07)}{0.28}\,I_{1994(4)} + \underset{(0.07)}{0.31}\,I_{1996(1)} - \underset{(0.07)}{0.50}\,I_{1996(3)} - \underset{(0.08)}{0.54}\,I_{1996(4)} \\ &+ \underset{(0.08)}{0.44}\,I_{1997(1)} + \underset{(0.08)}{0.40}\,I_{1997(2)} - \underset{(0.08)}{0.54}\,I_{1999(4)} + \underset{(0.08)}{0.37}\,I_{2000(1)} \end{aligned} \quad (4.16)$$

| | | |
|---|---|---|
| Vector SEM-AR 1-5 test: | F(80,365) | = 1.2321 [0.1043] |
| Vector ARCH 1-4 test: | F(64,409) | = 0.9374 [0.6142] |
| Vector Normality test: | $X^2(8)$ | = 3.5568 [0.8947] |

FIML, Sample: 1988(2) – 2019(4), N=127

Included above are the misspecification tests which were described earlier, only now they are in vector form to capture the full system. Again, the battery of tests shows no indications of any misspecification.

The multiple-equation model was estimated using an estimation method called Full Information Maximum Likelihood (FIML). FIML uses all available data to estimate the model parameters, including incomplete observations. It does so by maximizing the likelihood function of the observed data, considering all available information. An advantage of using FIML compared to OLS, is that FIML takes correlations between structural residuals into account.

The table below shows the correlations between the estimated residuals of the endogenous variables in the SEM. Most correlation terms are so close to zero that the estimation using OLS would have been almost identical to the estimation using FIML. However, particularly the correlation between $\Delta yf$ and $\Delta g$ is somewhat different from zero, which will be a source of different (but more efficient) estimation of the model compared to OLS. This can be seen in the equation system (4.13) – (4.16), which have slightly changed coefficients compared to the conditional and marginal models estimated separately as single equation models.

**Table 4.4:** Correlation of structural residuals in SEM (4.13) - (4.16)

| | $\Delta yf$ | $\Delta g$ | $\Delta emi$ | $\Delta joil$ |
|---|---|---|---|---|
| $\Delta yf$ | 0.0082328 | 0.15706 | -0.085298 | 0.062247 |
| $\Delta g$ | 0.15706 | 0.018724 | -0.015481 | 0.053512 |
| $\Delta emi$ | -0.085298 | -0.015481 | 0.010635 | -0.10188 |
| $\Delta joil$ | 0.062247 | 0.053512 | -0.10188 | 0.078413 |

### 4.3.4 Forecasting with the use of the Norwegian equation system

To create the economic scenario of no pandemic, dynamic forecasts in OxMetrics8 are used for the period 2020(1) to 2022(3). This forecast period is chosen because so-called in-sample forecasts can be performed, allowing for a comparison between our modelled counterfactual and the actual economic development. Dynamic forecasts differ from h-step forecasts, which assume that the underlying model remains constant over the forecast horizon, i.e., it estimates h steps (quarters) from a certain period. On the other hand, dynamic forecasts account for changes in the underlying relationships between the input variables and the output variable over time. This makes dynamic forecasts typically more accurate than h-step forecasts.

Dynamic forecast are multi-step forecasts obtained by simulation of a system of model equations (Nymoen, 2019). They are generated by recursively updating the model parameters over the forecast horizon, using the most recent observations of the input variables. This allows the model to adapt to changing patterns in the input variables, and to account for any shocks or unexpected events that may occur during the forecast period.

Before the forecasts, the following identity equations for $yf$ and $yf_{ECM}$ were added to the multiple-equation model. The inclusion of the identity equations is important for when the system is being forecasted. The equations yield simulated values for lagged values of $yf_t$ and $yf_{ECM_t}$ in the forecast period. The data for these variables would not have been available prior to the forecast period.

$$yf_t = 1\ \Delta yf_t + 1\ yf_{t-1} \tag{4.17}$$

$$yf_{ECM_t} = 1\ \Delta yf_t - 0.47\ \Delta g_t - 0.31\ \Delta emi_t + 1\ yf_{ECM_{t-1}} \tag{4.18}$$

Finally, the estimated forecasts for the period from 2020(1) to 2022(3) for the system of equations are plotted in figure 4.4 below. The forecast of $yf$ (bottom right) is what will be used as a counterfactual denoting a forecast with date of origin before the pandemic, i.e., prior to the shock. The error bands in the figure show the forecast errors, which naturally are increasing over time as more uncertainty is accounted for in the estimations. Because the forecast errors are asymptotically standard normal, they are estimated using standard t-values, which is common practice in the literature.

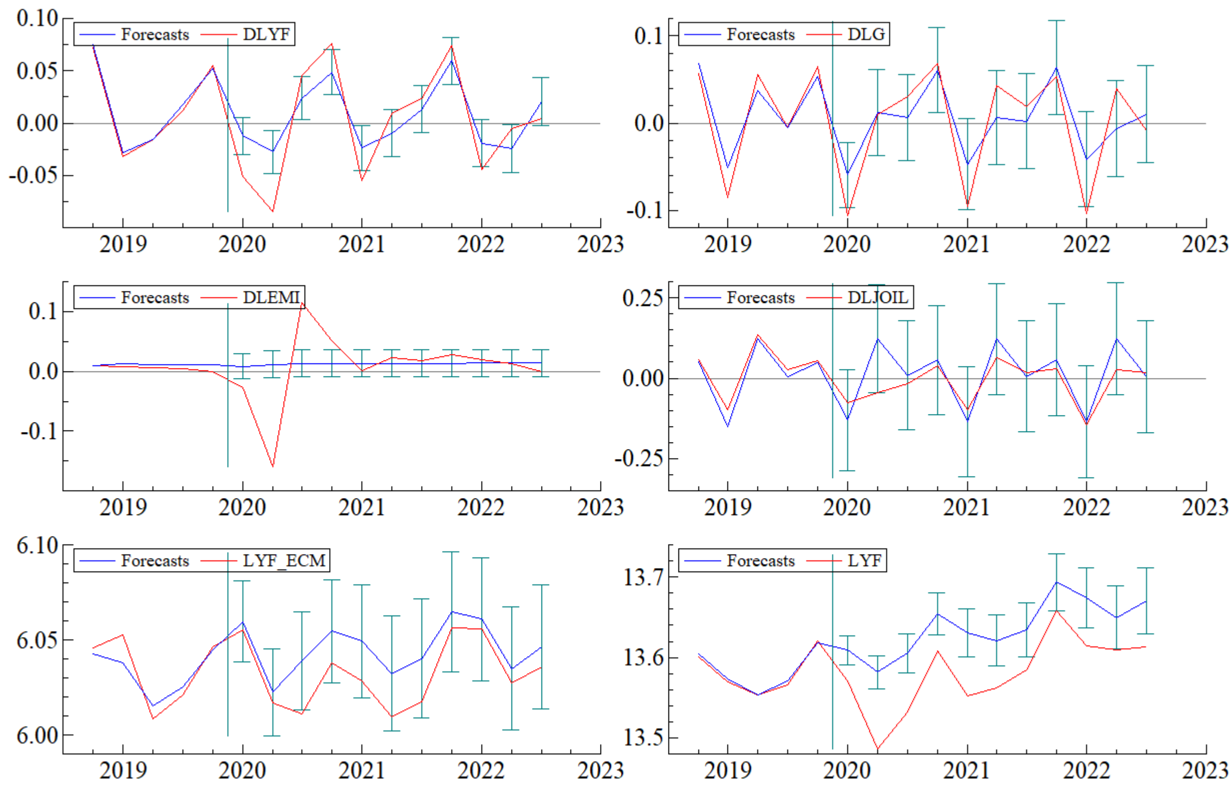


**Figure 4.4:** Forecasts for SEM (4.13) - (4.18), 2020(1) - 2022(3)

## 4.4 Aggregate model of the Swedish economy

In this chapter, the approach of building an aggregate model for the Swedish economy will be presented. The approach and the methods will in a large degree be similar to those presented in the previous chapter, when constructing the Norwegian model. As with the Norwegian GDP-equation, I began by estimating a Swedish GDP-equation with help from Autometrics to choose significant explanatory variables. Then, I tested for the same long-run relationship between GDP (Y), government expenditure (G) and foreign GDP growth (EMI). Subsequently, the GDP-equation was estimated with the ECM-variable included, followed by an estimation of the marginal model. Finally, I forecasted the whole multiple-equation model.

### 4.4.1 Specifying an equation for Swedish GDP using Autometrics

The variables that were included in the GUM for the estimation of Swedish GDP are presented in the following. x lags of the dependent variable (e.g., $\Delta y_{t-1}, \dots, \Delta y_{t-x}$) were included, as the lagged values are likely to be explanatory for current GDP growth, in Sweden as it was in Norway. Lagged level variables: GDP ($y_{t-1}$), government expenditure ($g_{t-1}$) and foreign trade-weighted GDP ($emi_{t-1}$) proxying as a market growth indicator, were also included. Note that EMI is the same variable as the constructed YF-variable in the re-estimated GDP-equation in the last chapter, which has been renamed for notational convenience. Also included were differenced variables of government expenditure ($\Delta g_t$), the market growth indicator ($\Delta emi_t$) and real interest rate (RLINF) with four lags each. Additionally, a variable for the number of working days in Sweden per quarter (*arbdag*) and centered trend seasonals ($CS_t$) were included. Finally, the significant variable proxying the difference in GDP-growth in Sweden and its trade partners ($(y - emi)_{t-1}$) from the re-estimation of NAM-2009, as well as impulse indicator variables for each observation in the sample were added to the GUM, by using the IIS method in Autometrics.

Choosing target size of 1 percent as overall significance level, the final model equation estimated using Autometrics was:

$$\begin{aligned}\Delta y_t = &\underset{(0.59)}{2.81}\, Constant - \underset{(0.02)}{0.14}\, y_{t-1} + \underset{(0.01)}{0.04}\, g_{t-1} + \underset{(0.009)}{0.05}\, emi_{t-1} + \underset{(0.03)}{0.14}\, \Delta emi \\ &+\underset{(0.02)}{0.08}\, \Delta emi_{t-1} + \underset{(0.02)}{0.09}\, \Delta emi_{t-2} - \underset{(0.01)}{0.04}\, \Delta g_{t-1} - \underset{(0.007)}{0.02}\, \Delta g_{t-2} \\ &- \underset{(0.004)}{0.02}\, \Delta g_{t-3} - \underset{(0.006)}{0.023} I_{2008(4)} + \underset{(0.006)}{0.018}\, I_{2011(3)} - \underset{(0.006)}{0.021}\, I_{2011(4)}\end{aligned} \tag{4.19}$$

| | | |
|---|---|---|
| OLS | Sample: 1996(3) - 2019(4) | N = 94 |
| $AR_{1-5}$: | F(5,76) = 1.4467 [0.2174] | $\hat{\sigma}100 = 0.59$ |
| $ARCH_{1-4}$: | F(4,86) = 1.6560 [0.1677] | RSS = 0.0029 |
| Normality: | $\chi^2(2)$ = 1.1566 [0.5609] | |

Autometrics retained only three indicator variables: The fourth quarter in 2008, the financial crisis, and the two last quarters of 2011, with opposite signed coefficients interpretable as a growth rate shocks. Note in particular that the lagged level variables $y_{t-1}$, $g_{t-1}$ and $emi_{t-1}$ were retained by the machine learning algorithm.

The t-value of the retained variable $y_{t-1}$ was -5.19. The t-value of -5.19 rejects the null hypothesis of no cointegration against the 1 percent critical value of the ECM-test with three I(1) variables, which is –4.09 (Ericsson & MacKinnon, 2002, p. 304). Based on this result I constructed the following ECM-variable:

$$y_{ECM_t} = y_t - \frac{\hat{\gamma}_1}{\hat{\pi}} g_t - \frac{\hat{\gamma}_2}{\hat{\pi}} emi_t \tag{4.20}$$

Where, as in the Norwegian GDP-equation, the "hat-notation" is used to represent estimated coefficients from the GDP-equation (4.19). Inserting for the estimated coefficients yielded the following ECM-equation of the long-run residuals:

$$y_{ECM_t} = y_t - 0.31\, g_t - 0.33\, emi_t \tag{4.21}$$

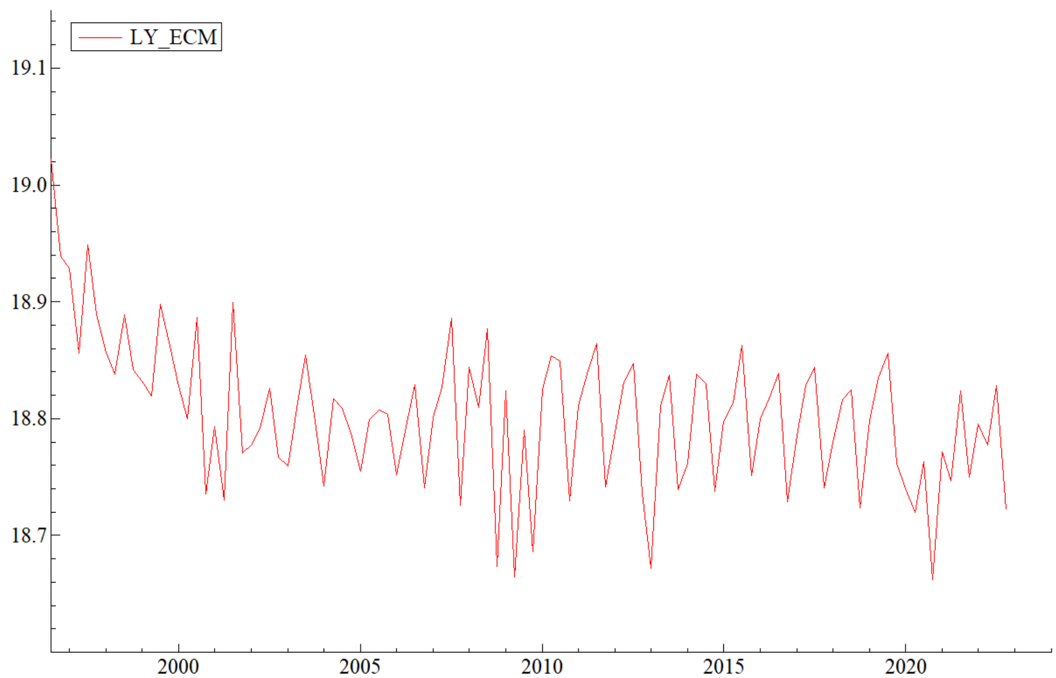


**Figure 4.5:** Plot of the $y_{ECM}$-variable defined by equation (4.21)

Sample: 1988(2) – 2022(3)

In the same was as for Norway, I did a robustness check of cointegration by using the Engle-Granger test. Like before, the test was performed on the static equation of the long-run relationship between $y_{t-1}$, $g_{t-1}$ and $emi_{t-1}$. The estimated equation became:

$$y_t = 0.08\ g_t + 0.33\ emi_t + 24.8 \tag{4.22}$$

Taking the estimated coefficients as given, the long-run residuals can be defined as the variable $y_{static}$:

$$y_{static_t} = y_t - 0.08\ g_t - 0.33\ emi_t \tag{4.23}$$

This static regression can be used to test the null hypothesis of no cointegration using an ADF-test. The critical values depend on the number of variables included in the "cointegration regression," which in this case is three. First, I ran the unit-root test, and compared the t-values estimated against the critical values in table 1 of MacKinnon (1990). Based on the results listed in the table below, the null hypothesis of no cointegration can be rejected at the 5% significance level. The p-value (i.e., t-prob) of the second differenced lag of the $y_{static}$-variable is significant, meaning that this the is the smallest model that would have given valid test results.

**Table 4.5:** ADF-test of $y_{static}$

Critical values from MacKinnon (1990): 5% = -3.74, 1% = -4.29 (N = 94, Constant)
Sample: 1996(3) – 2019(4)

| D-lag | t-adf | t-DY_lag | t-prob |
|---|---|---|---|
| 5 | -2.820 | 1.692 | 0.0942 |
| 4 | -2.931 | 6.188 | 0.0000 |
| 3 | -5.309** | -2.861 | 0.0053 |
| 2 | -4.505** | -1.009 | 0.3158 |
| 1 | -4.390** | -3.740 | 0.0003 |
| 0 | -4.156* | | |

The use and interpretation of the ADF-test is in this case more complicated due to the residual autocorrelation first being strongly negative, followed by being positive. This is because the negative autocorrelation for the first residuals in the sample leads to an overestimation of the estimated standard deviations. Consequently, the result is not as strong as in the Norwegian case, but still strong enough to support the hypothesis of a long-run relationship between the three variables. Furthermore, the ECM-test has already given a significant rejection of the null hypothesis. The ECM-test and EG-test are similar in terms of Type-I error probabilities, but the ECM-test generally has lower probability of Type-II errors (Nymoen, 2019, p. 354). As such, the power of the ECM-test is better and is naturally given more weight in the literature.

Proceeding with the $y_{ECM}$-variable included in the GUM, replacing the trend variables $y_{t-1}$, $g_{t-1}$ and $emi_{t-1}$, I now have a model with stationary variables. Rerunning Autometrics with the updated GUM yielded the following equation:

$$\begin{aligned}\Delta y_t =& \underset{(0.51)}{2.85}\, Constant - \underset{(0.02)}{0.15}\, y_{ECM_{t-1}} + \underset{(0.03)}{0.14}\, \Delta emi + \underset{(0.02)}{0.07}\, \Delta emi_{t-1} \\ &+ \underset{(0.02)}{0.09}\, \Delta emi_{t-2} - \underset{(0.006)}{0.04}\, \Delta g_{t-1} - \underset{(0.005)}{0.02}\, \Delta g_{t-2} - \underset{(0.003)}{0.02}\, \Delta g_{t-3} \\ &- \underset{(0.006)}{0.024} I_{2008(4)} + \underset{(0.006)}{0.018}\, I_{2011(3)} - \underset{(0.006)}{0.021}\, I_{2011(4)} \end{aligned} \tag{4.24}$$

| | | |
|---|---|---|
| OLS | Sample: 1996(3) - 2019(4) | N = 94 |
| $AR_{1-5}$: | F(5,78) = 1.4634 [0.2114] | $\hat{\sigma}100 = 0.59$ |
| $ARCH_{1-4}$: | F(4,86) = 1.6067 [0.1799] | RSS = 0.0029 |
| Normality: | $\chi^2(2)$ = 1.1125 [0.5734] | |

Except for the $y_{ECM}$-variable replacing the trend variables, the estimated model equation (4.24) consists of the same variables as (4.19). Therefore, the RSS and $\hat{\sigma}100$ are practically identical to equation (4.19). For the same reason are the misspecification tests mostly unchanged, still showing no signs of misspecification of the model equation.

Figures 4.6 and 4.7 show the goodness of fit plots of $\Delta y$ and parameter constancy tests of $y_{ECM_{t-1}}$, respectively. The recursive plot shows the estimated single equation model of $\Delta y$ graphically. The set of parameter constancy tests are graphically showing the parameter constancy of the ECM-variable, which again are showing no significant issues with parameter constancy. The plots are showing the estimated coefficient (with estimated standard errors), the 1-step residuals, 1-step Chow tests and break-point Chow tests of $y_{ECM_{t-1}}$.

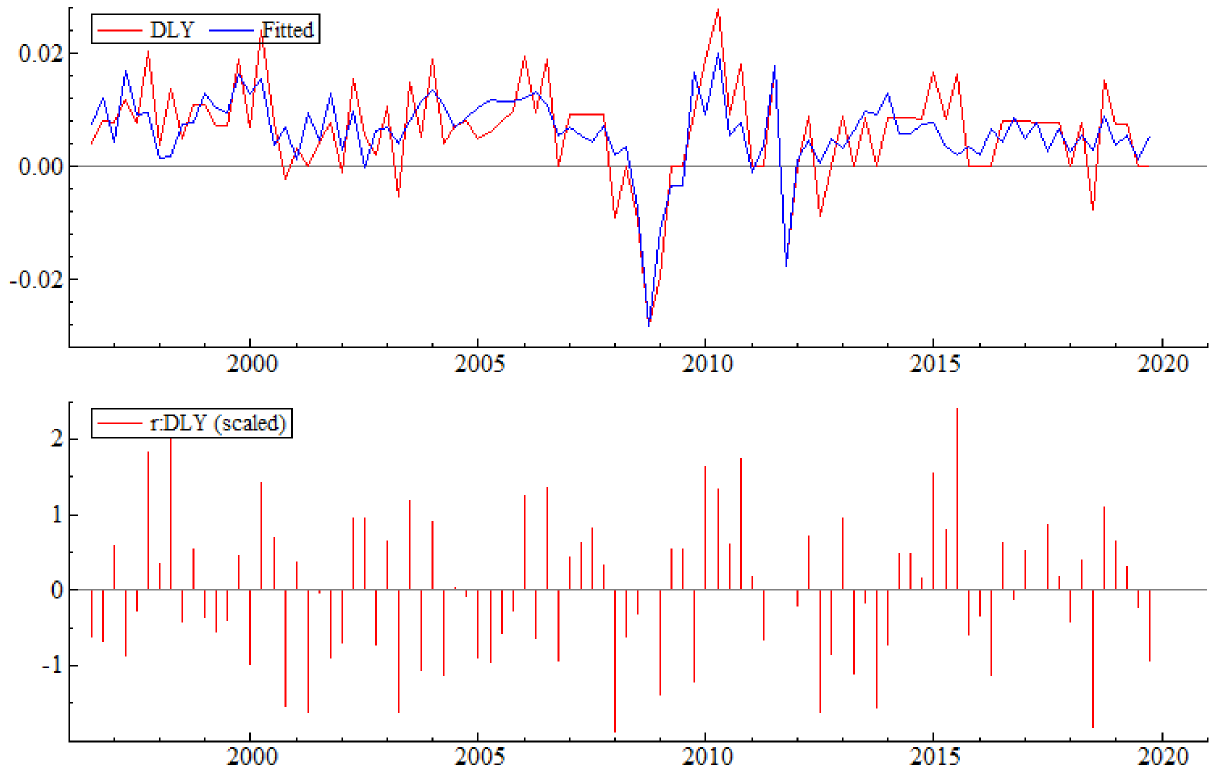


**Figure 4.6:** Actual and fitted values of $\Delta y$ (DLY in figure) in equation (4.24), and scaled residuals (DLY scaled) in lower panel

Upper left panel of figure 4.7, which is showing the estimated coefficient of $y_{ECM_{t-1}}$, shows that there is some recursive instability in the estimated coefficient before 2010. This indicates stronger error correction after 2010 relative to before. However, it is always inside the 95% confidence interval, and moves towards the long-run solution, indicating that the parameter constancy is upheld overall.

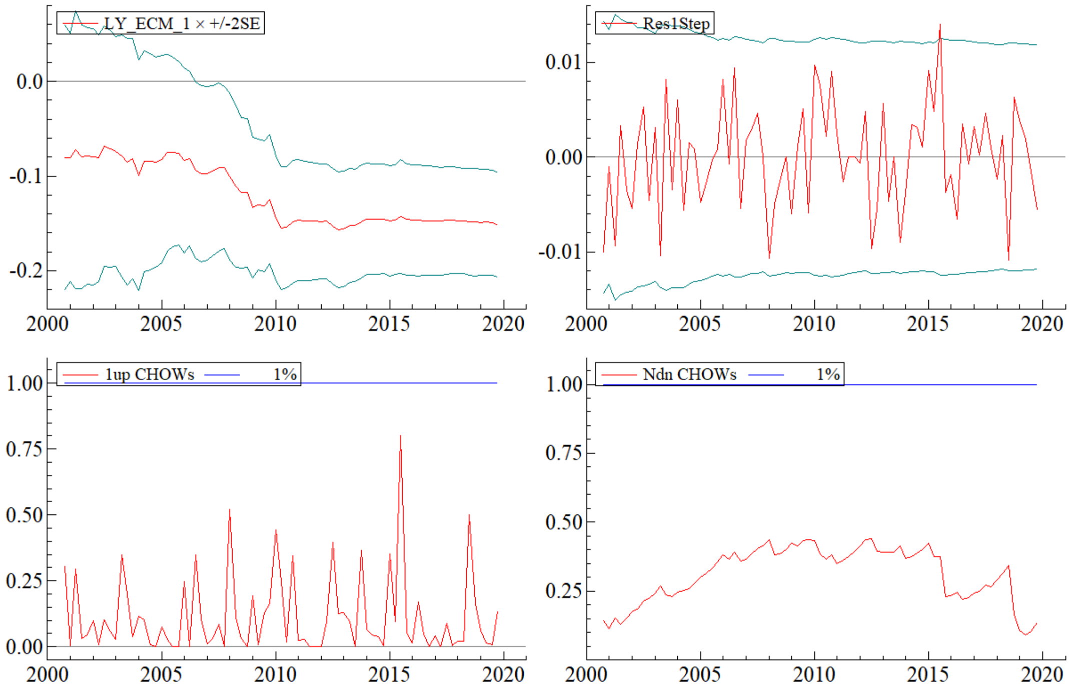


**Figure 4.7:** Recursive parameter constancy tests for $y_{ECM_{t-1}}$ in equation (4.24)

### 4.4.2 The estimated marginal equations in the Swedish equation system

I chose to continue with the GDP-equation containing the ECM-variable (4.24) as the conditional model in the equation system. Consequently, I needed to estimate the marginal models for the remaining endogenous variables in the system. The marginal models were specified in the same manner as for the Norwegian model: Using only lags of the dependent variable, deterministic variables, and impulse indicators, i.e., they will be autoregressive processes. In equation (4.24), there are two endogenous variables: $\Delta g$ and $\Delta emi$, i.e., government expenditure and the foreign market indicator on log difference form.

I started with modelling $\Delta g$, and included six lags of the dependent variable, as well as the deterministic variable $arbdag$ and centered quarterly seasonal dummies in the GUM. Once more allowing for IIS, impulse indicator for each period in the sample were also added. This process was then repeated for $\Delta emi$, only changing the dependent variable (and its lags). Because both variables are economic indicators with long-term trends, constant terms were fixed when estimating. The estimated coefficients of the retained variables in the marginal models are given in table 4.6 below.

| **Table 4.6:** Estimated coefficients in the marginal models of Swedish model | | | |
|---|---|---|---|
| $\Delta g$ | | $\Delta emi$ | |
| Variable name | Coefficient | Variable name | Coefficient |
| $Constant$ | $\underset{(0.008)}{0.005}$ | $Constant$ | $\underset{(0.001)}{0.008}$ |
| $\Delta g_{t-1}$ | $\underset{(0.05)}{-0.83}$ | $\Delta emi_{t-1}$ | $\underset{(0.04)}{0.27}$ |
| $\Delta g_{t-2}$ | $\underset{(0.06)}{-0.65}$ | $\Delta emi_{t-2}$ | $\underset{(0.04)}{0.15}$ |
| $\Delta g_{t-3}$ | $\underset{(0.05)}{-0.75}$ | $I_{1997(1)}$ | $\underset{(0.01)}{0.08}$ |
| $\Delta g_{t-6}$ | $\underset{(0.05)}{-0.18}$ | $I_{1997(2)}$ | $\underset{(0.01)}{0.08}$ |
| $\Delta arbdag_{t-1}$ | $\underset{(0.17)}{0.64}$ | $I_{1997(3)}$ | $\underset{(0.01)}{-0.04}$ |
| $I_{2000(4)}$ | $\underset{(0.07)}{0.27}$ | $I_{1998(1)}$ | $\underset{(0.01)}{0.06}$ |
| $I_{2009(2)}$ | $\underset{(0.07)}{0.42}$ | $I_{1998(2)}$ | $\underset{(0.01)}{-0.04}$ |
| $I_{2010(2)}$ | $\underset{(0.08)}{-0.32}$ | $I_{1998(3)}$ | $\underset{(0.01)}{0.07}$ |
| $I_{2013(1)}$ | $\underset{(0.07)}{0.39}$ | $I_{2001(1)}$ | $\underset{(0.01)}{-0.03}$ |
| | | $I_{2003(4)}$ | $\underset{(0.01)}{0.03}$ |
| | | $I_{2006(3)}$ | $\underset{(0.01)}{-0.03}$ |
| | | $I_{2008(4)}$ | $\underset{(0.01)}{-0.05}$ |
| AR 1-5 test: | | AR 1-5 test: | |
| F(5,79) = 1.9751 [0.0914] | | F(5,76) = 1.4178 [0.2275] | |
| ARCH 1-4 test: | | ARCH 1-4 test: | |
| F(4,86) = 0.9539 [0.4371] | | F(4,84) = 1.1103 [0.3570] | |
| Normality test: | | Normality test: | |
| $X^2(2)$ = 0.1531 [0.9263] | | $X^2(2)$ = 8.5976 [0.0136]* | |

OLS, Sample: 1996(3) – 2019(4), N=94

### 4.4.3 The multiple-equation model for Swedish GDP

The complete system was put together by estimating a SEM with $\Delta y$ as the conditional model and $\Delta g$ and $\Delta emi$ as marginal models. In the following, the whole equation system is presented, where (4.25) is the conditional model of the system and (4.26) – (4.27) are the marginal models.

$$\begin{aligned}\Delta y_t = &\underset{(0.51)}{2.85}\, Constant - \underset{(0.02)}{0.15}\, y_{ECM_{t-1}} + \underset{(0.03)}{0.14}\, \Delta emi + \underset{(0.02)}{0.07}\, \Delta emi_{t-1} \\ &+ \underset{(0.02)}{0.09}\, \Delta emi_{t-2} - \underset{(0.006)}{0.04}\, \Delta g_{t-1} - \underset{(0.005)}{0.02}\, \Delta g_{t-2} - \underset{(0.003)}{0.02}\, \Delta g_{t-3} \\ &- \underset{(0.006)}{0.024} I_{2008(4)} + \underset{(0.006)}{0.018}\, I_{2011(3)} - \underset{(0.006)}{0.021}\, I_{2011(4)}\end{aligned} \tag{4.25}$$

$$\begin{aligned}\Delta g_t = &\underset{(0.008)}{0.005}\, Constant - \underset{(0.05)}{0.81}\, \Delta g_{t-1} - \underset{(0.06)}{0.63}\, \Delta g_{t-2} - \underset{(0.04)}{0.74}\, \Delta g_{t-3} - \underset{(0.05)}{0.19}\, \Delta g_{t-6} \\ &+ \underset{(0.17)}{0.69}\, \Delta arbdag_{t-1} + \underset{(0.07)}{0.32}\, I_{2000(4)} + \underset{(0.07)}{0.40}\, I_{2009(2)} - \underset{(0.08)}{0.34}\, I_{2010(2)} + \underset{(0.07)}{0.38}\, I_{2013(1)}\end{aligned} \tag{4.26}$$

$$\begin{aligned}\Delta emi_t = &\underset{(0.001)}{0.008}\, Constant + \underset{(0.04)}{0.27}\, \Delta emi_{t-1} + \underset{(0.04)}{0.15}\, \Delta emi_{t-2} + \underset{(0.01)}{0.08}\, I_{1997(1)} \\ &+ \underset{(0.01)}{0.09}\, I_{1997(2)} - \underset{(0.01)}{0.04}\, I_{1997(3)} + \underset{(0.01)}{0.06}\, I_{1998(1)} - \underset{(0.01)}{0.05}\, I_{1998(2)} + \underset{(0.01)}{0.07}\, I_{1998(3)} \\ &- \underset{(0.01)}{0.03}\, I_{2001(1)} + \underset{(0.01)}{0.03}\, I_{2003(4)} - \underset{(0.01)}{0.03}\, I_{2006(3)} - \underset{(0.01)}{0.04}\, I_{2008(4)}\end{aligned} \tag{4.27}$$

| | | |
|---|---|---|
| Vector SEM-AR 1-5 test: | F(45,196) | = 1.0176 [0.4509] |
| Vector ARCH 1-4 test: | F(36,225) | = 1.0557 [0.3913] |
| Vector Normality test: | $X^2(6)$ | = 8.6843 [0.1921] |

FIML, Sample: 1996(3) – 2019(4), N=94

Note that the vector versions of the misspecification tests reported below the equations shows no indications of any misspecification of the system. Also included in table 4.7 below are the estimated correlations between the estimated residuals between the endogenous variables. Note that the correlations between some of the residuals are a bit larger compared to the Norwegian case (table 4.4). Especially the correlation between the residuals of $\Delta y$ and $\Delta emi$ is slightly different from zero. Consequently, estimating the system using FIML will lead to some changes compared to the OLS estimations of the single equations in the system. This is why the estimated coefficients in SEM (4.25) – (4.27) are somewhat different from the coefficients in the single equation estimations.

**Table 4.7:** Correlation of structural residuals in SEM (4.25) - (4.27)

| | $\Delta y$ | $\Delta g$ | $\Delta emi$ |
|---|---|---|---|
| $\Delta y$ | 0.0061382 | 0.19160 | 0.38038 |
| $\Delta g$ | 0.19160 | 0.076488 | -0.088388 |
| $\Delta emi$ | 0.38038 | -0.088388 | 0.013727 |

### 4.4.4 Forecasting with the use of the Swedish equation system

Finally, following the same steps as with the Norwegian system, I used dynamic forecasts to create the scenario of no pandemic in the period 2020(1) to 2022(3). The below identity equations for $y$ and $y_{ECM}$ were included in the forecasting model:

$$y_t = 1\ \Delta y_t + 1\ y_{t-1} \tag{4.28}$$

$$y_{ECM_t} = 1\ \Delta y_t - 0.31\ \Delta g_t - 0.33\ \Delta emi_t + 1\ y_{ECM_{t-1}} \tag{4.29}$$

The estimated forecasts from the period from 2020(1) to 2022(3) are plotted in figure 4.8. Note especially the forecast of $y$ (notation LY in figure), which will be the counterfactual, i.e., the estimated effect of no-covid in the Swedish economy. This counterfactual denotes a forecast with date of origin prior to the pandemic. As before, the error bands show the forecast errors, which were estimated using standard t-values.


**Figure 4.8:** Forecasts for SEM (4.25) - (4.29), 2020(1) - 2022(3)

## 4.5 Final form equations

To have another comparable estimation of GDP for Mainland-Norway and Sweden, I chose to also include estimates using final form equations. To estimate the final form equations for GDP in Mainland-Norway and Sweden, I used the machine learning algorithm Autometrics once again. The estimation samples were the same as for the respective aggregate models. The respective GUMs consisted of lagged differences of the dependent variable, centered seasonal dummies and the impulse indicator dummies.

Running Autometrics with target size = 0.01 and IIS, the algorithm chose to retain $\Delta yf_{t-1}$ and $\Delta yf_{t-12}$ when estimating the GDP for Mainland-Norway. Moreover, all three seasonal dummies were retained, as well as an impulse indicator for the second quarter of 1997. The final equation became equation (4.30), with information on model fit and misspecification tests included below. There was no indication of misspecification in the test results.

$$\Delta yf_t = \underset{(0.0016)}{0.006}\ Constant - \underset{(0.07)}{0.46}\ \Delta yf_{t-1} + \underset{(0.06)}{0.20}\ \Delta yf_{t-12} - \underset{(0.008)}{0.043}\ CS_t - \underset{(0.006)}{0.068}\ CS_{t-1} - \underset{(0.006)}{0.05}\ CS_{t-2} + \underset{(0.016)}{0.06}\ I_{1997(2)} \tag{4.30}$$

| | | |
|---|---|---|
| OLS | Sample: 1988(2) - 2019(4) | N = 94 |
| $AR_{1-5}$: | F(5,115) = 2.0542 [0.0762] | $\hat{\sigma}100 = 0.61$ |
| $ARCH_{1-4}$: | F(4,119) = 0.87817 [0.4793] | RSS = 0.0033 |
| Normality: | $\chi^2(2)$ = 0.39750 [0.8198] | |

Running Autometrics with target size 0.01 and IIS, this time on the Swedish data, the algorithm retained only one lagged dependent variable, $\Delta y_{t-2}$. In addition, it chose to retain seven impulse dummies. The dummies cover quarter with unusually large development: Three quarters during the financial crisis, and selected quarters in the following years. The misspecification tests to the final equation (4.31) showed no signs of model misspecification.

$$\Delta y_t = \underset{(0.0008)}{0.006}\ Constant + \underset{(0.07)}{0.17}\ \Delta y_{t-2} - \underset{(0.006)}{0.016}\ I_{2008(1)} - \underset{(0.006)}{0.013}\ I_{2008(3)} - \underset{(0.006)}{0.034}\ I_{2008(4)} - \underset{(0.006)}{0.023}\ I_{2009(1)} + \underset{(0.006)}{0.019}\ I_{2010(2)} - \underset{(0.006)}{0.023}\ I_{2011(4)} - \underset{(0.006)}{0.015}\ I_{2012(3)} \tag{4.31}$$

| | | |
|---|---|---|
| OLS | Sample: 1996(3) - 2019(4) | N = 94 |
| $AR_{1-5}$: | F(5,80) = 0.93522 [0.4628] | $\hat{\sigma}100 = 0.61$ |
| $ARCH_{1-4}$: | F(4,86) = 1.6618 [0.1663] | RSS = 0.0032 |
| Normality: | $\chi^2(2)$ = 0.51131 [0.7744] | |

# Chapter 5: Results about the impact of the Covid-19 pandemic on GDP

In this chapter, results from the aggregate model estimations are presented. Using the aggregate models and the final form equations from the previous chapter, the estimated differences between the actual and counterfactual for the two countries are presented as well. Next, these differences will be compared between Norway and Sweden. Finally, there will be a discussion on the findings, their credibility, and the findings in a broader picture.

## 5.1 Difference between actuals and counterfactuals in Norway

To estimate the economic impact of Covid-19 and the policy responses to it, I use the differences between the forecasted and actual values of Mainland-Norway GDP, i.e., the measure corresponding to $Diff_{II}\boldsymbol{y}_t$, in chapter 2.2. The results in figure 5.1 can be interpreted as the estimated difference between my model's estimated GDP given no pandemic and the actual GDP numbers. The plotted values are given in table 5.1.

For the Norwegian economy, figure 5.1 shows that the main finding is a long-Covid loss in GDP growth. For the whole forecast period, there is an estimated loss in GDP due to the pandemic and the policy response. The most major estimated difference can be found in the second quarter of 2020, when a strict lockdown was in full effect in Norway. Moreover, when a strong policy response once more was implemented in the first quarter of 2021 to deal with the Delta variant, the economy suffered another draw-down. The decline in GDP growth was not as strong as during the first wave. Reasons could be plentiful, but most importantly, the policy response was now more efficiently directed, and the economy had adapted to handle the situation better (Andersen et al., 2022). In both cases, despite growth in the following quarters, the economy did not catch up to the forecasted counterfactual GDP-level.

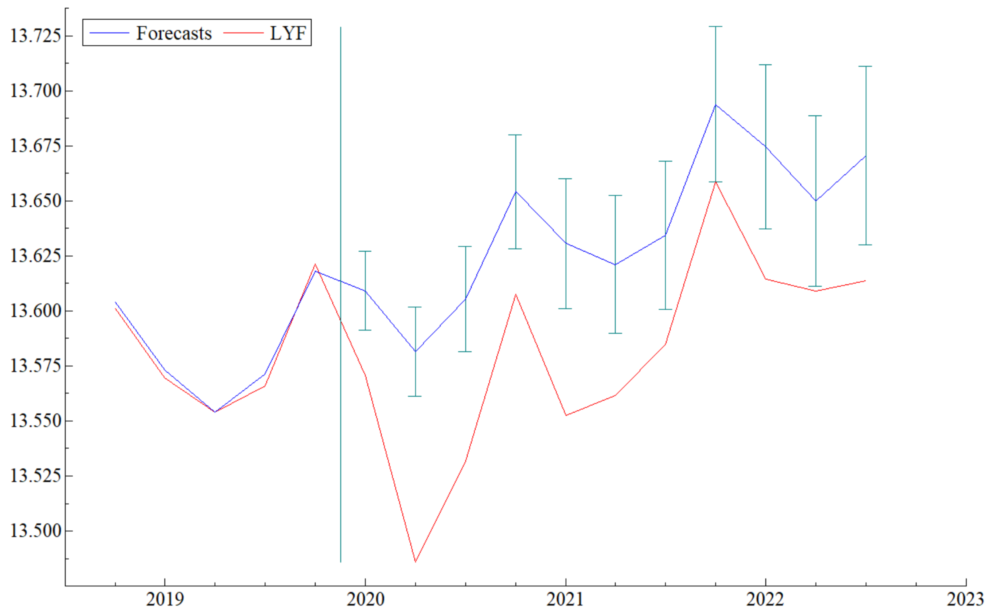


**Figure 5.1:** Difference between actuals and counterfactuals in multiple-equation model for Mainland-Norway GDP

**Table 5.1:** Log of GDP Mainland-Norway. Results using the aggregate model for Mainland-Norway.

| (1) Horizon | (2) Counterfactual | (3) Actual | (4) Difference (2) – (3) | (5) t-value | (6) Difference Percent |
|---|---|---|---|---|---|
| 2020(1) | 13.6091 | 13.5706 | 0.038476 | 4.306 | 3.8 |
| 2020(2) | 13.5815 | 13.4859 | 0.095610 | 9.444 | 9.5 |
| 2020(3) | 13.6053 | 13.5319 | 0.073373 | 6.130 | 7.3 |
| 2020(4) | 13.6541 | 13.6076 | 0.046491 | 3.583 | 4.6 |
| 2021(1) | 13.6306 | 13.5526 | 0.078033 | 5.278 | 7.8 |
| 2021(2) | 13.6211 | 13.5617 | 0.059387 | 3.785 | 5.9 |
| 2021(3) | 13.6343 | 13.5847 | 0.049627 | 2.949 | 4.9 |
| 2021(4) | 13.6937 | 13.6585 | 0.035247 | 1.997 | 3.5 |
| 2022(1) | 13.6745 | 13.6145 | 0.059996 | 3.217 | 5.9 |
| 2022(2) | 13.6499 | 13.6091 | 0.040777 | 2.100 | 4.0 |
| 2022(3) | 13.6705 | 13.6136 | 0.056907 | 2.809 | 5.6 |

The results from the final form equation estimation of GDP in Mainland-Norway are quite close to those obtained from the forecasts of the multiple-equation model. However, in general, the estimated differences in table 5.2 below are slightly smaller, especially towards the end of the forecast period. The last period with a significant difference-from-counterfactual (at the 5% significance level) is the first quarter of 2021. For later quarters, the confidence intervals include the actual GDP numbers. Hence the method of final form equation do not yield a significant long-Covid effect on Mainland-Norway GDP.

**Table 5.2:** Log of GDP Mainland-Norway. Results using final form equation.

| (1) Horizon | (2) Counterfactual | (3) Actual | (4) Difference (2) – (3) | (5) t-value | (6) Difference Percent |
|---|---|---|---|---|---|
| 2020(1) | 13.5969 | 13.5706 | 0.026336 | 1.585 | 2.6 |
| 2020(2) | 13.5806 | 13.4859 | 0.094638 | 5.015 | 9.4 |
| 2020(3) | 13.5863 | 13.5319 | 0.054441 | 2.406 | 5.4 |
| 2020(4) | 13.6457 | 13.6076 | 0.038106 | 1.518 | 3.8 |
| 2021(1) | 13.6129 | 13.5526 | 0.060344 | 2.182 | 6.0 |
| 2021(2) | 13.6101 | 13.5617 | 0.048472 | 1.624 | 4.8 |
| 2021(3) | 13.6039 | 13.5847 | 0.019141 | 0.599 | 1.9 |
| 2021(4) | 13.6695 | 13.6585 | 0.010969 | 0.323 | 1.0 |
| 2022(1) | 13.6367 | 13.6145 | 0.022229 | 0.621 | 2.2 |
| 2022(2) | 13.6277 | 13.6091 | 0.018506 | 0.493 | 1.8 |
| 2022(3) | 13.6315 | 13.6136 | 0.017894 | 0.456 | 1.7 |

## 5.2 Difference between actuals and counterfactuals in Sweden

In figure 5.2 below, the difference between actuals and counterfactuals for Sweden are presented. The plotted values are given in table 5.3, where, due to the definition of difference-from-counterfactual, negative values imply higher growth in GDP relative to forecasted growth. The results indicate that Sweden only suffered one large downturn in the economy in the second quarter of 2020, followed by a quick catch-up the following quarter, almost covering the decline in the quarter before. Moreover, throughout 2021, the complete difference between the forecasted values from the aggregate model and the actual GDP values, has been covered. Finally, the first three quarters of 2022 were characterized by no GDP-growth, resulting in no difference between the forecasted values and actuals as per the third quarter of 2022. The results indicate no long-run effect from the pandemic.

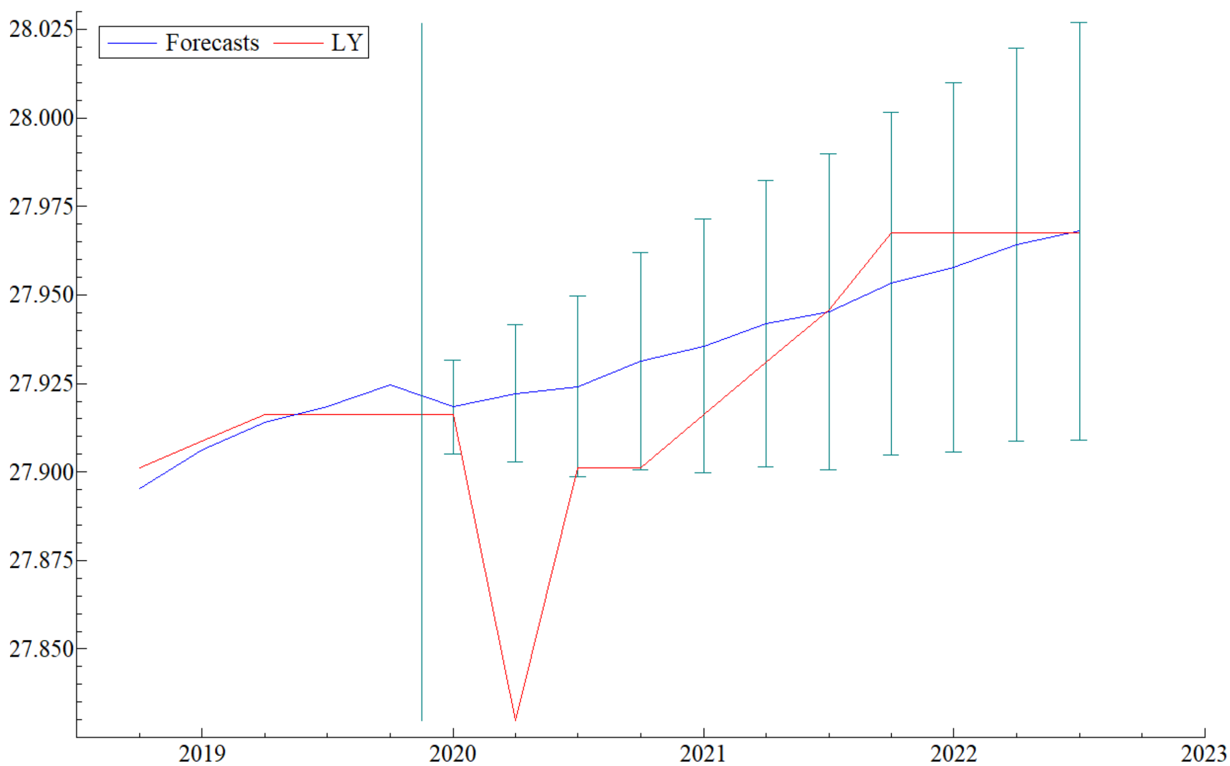


**Figure 5.2:** Difference between actuals and counterfactuals in multiple-equation model for Sweden

**Table 5.3:** Log of GDP Sweden. Results using the aggregate model for Sweden.

| (1) Horizon | (2) Counterfactual | (3) Actual | (4) Difference (2) – (3) | (5) t-value | (6) Difference Percent |
|---|---|---|---|---|---|
| 2020(1) | 27.9184 | 27.9162 | 0.0022278 | 0.336 | 0.2 |
| 2020(2) | 27.9222 | 27.8299 | 0.092280 | 9.547 | 9.2 |
| 2020(3) | 27.9241 | 27.9010 | 0.023078 | 1.815 | 2.3 |
| 2020(4) | 27.9313 | 27.9010 | 0.030273 | 1.969 | 3.0 |
| 2021(1) | 27.9355 | 27.9162 | 0.019349 | 1.079 | 1.9 |
| 2021(2) | 27.9418 | 27.9311 | 0.010699 | 0.530 | 1.0 |
| 2021(3) | 27.9452 | 27.9458 | -0.00067264 | -0.030 | 0.0 |
| 2021(4) | 27.9532 | 27.9675 | -0.014251 | -0.589 | -1.4 |
| 2022(1) | 27.9579 | 27.9675 | -0.00952758 | -0.367 | -0.9 |
| 2022(2) | 27.6942 | 27.9675 | -0.0032920 | -0.118 | -0.3 |
| 2022(3) | 27.9680 | 27.9675 | 0.00049721 | 0.017 | 0.0 |

The results from the final form equation estimation of Swedish GDP were quite close to the forecast the aggregate model yielded. In this case, the final form equation estimated a more persistent effect of Covid-19. While the last significant t-value is for the forecast in the first quarter of 2021, the succeeding (insignificant) differences are all positive, indicating that when using this estimation method, the economy does not fully catch up the loss from the pandemic-shock. Due to the large uncertainty in forecasting, the indications are not significant, and there is not found a significant long-Covid effect on GDP in Sweden. The results are listed in table 5.4 below.

**Table 5.4:** Log of GDP Sweden. Results using final form equation.

| (1) Horizon | (2) Counterfactual | (3) Actual | (4) Difference (2) – (3) | (5) t-value | (6) Difference Percent |
|---|---|---|---|---|---|
| 2020(1) | 27.9224 | 27.9162 | 0.0062493 | 1.010 | 0.6 |
| 2020(2) | 27.9287 | 27.8299 | 0.098827 | 11.299 | 9.8 |
| 2020(3) | 27.9360 | 27.9010 | 0.034997 | 3.077 | 3.4 |
| 2020(4) | 27.9434 | 27.9010 | 0.042343 | 3.137 | 4.2 |
| 2021(1) | 27.9509 | 27.9162 | 0.034730 | 2.252 | 3.4 |
| 2021(2) | 27.9585 | 27.9311 | 0.027343 | 1.596 | 2.7 |
| 2021(3) | 27.9660 | 27.9458 | 0.020209 | 1.081 | 2.0 |
| 2021(4) | 27.9736 | 27.6975 | 0.0061206 | 0.304 | 0.6 |
| 2022(1) | 27.9812 | 27.6975 | 0.013699 | 0.637 | 1.3 |
| 2022(2) | 27.9888 | 27.6975 | 0.021278 | 0.935 | 2.1 |
| 2022(3) | 27.9964 | 27.6975 | 0.028857 | 1.204 | 2.8 |

## 5.3 Comparison of results between Norway and Sweden

The results obtained using the model framework described in chapter 4, are culminating in the comparison of the results between Norway and Sweden. The estimated differences-from-counterfactuals for the aggregate models are listed in table 5.6. By doing a literature search, comparable studies have also measured the economic effect of Covid-19 in Norway. The estimates from Bjertnæs et al. (2021) and von Brasch et al. (2022) are also listed in table 5.6 as comparable results. These studies used a combination of $Diff_I \boldsymbol{y}_t$ and $Diff_{II} \boldsymbol{y}_t$ described in chapter 2.2 and in Nymoen (2023).

For Sweden, the literature search yielded no directly comparable results. However, to obtain a comparable result also for the Swedish economy, I found a forecast from December 2019 provided by Konjunkturinstitutet (The National Institute of Economic Research) (2019) as well as a forecast from January 2020 made by the Swedish Ministry of Finance (Regeringskansliet, 2020). Because the two forecasts are practically identical, only the forecast from

Konjunkturinstitutet have been included. However, the fact that the two forecasts are so alike, strengthens the credibility of Konjunkurinstitutet's forecast, indicating that this is the consensus prior to the shock occurring.

The forecasted values were collected in percentage changes per quarter. As such, to get a comparable result, I used the value of GDP in Sweden in the fourth quarter of 2019 as a starting point and then used the forecasted growth values to project GDP-values in the eleven succeeding quarters. The forecasted values, as well as the differences-from-counterfactuals, are listed in table 5.5 below.

**Table 5.5:** Log of GDP Sweden. Results using Konjunkturinstitutet's forecast from December 2019.

| (1) Horizon | (2) Counterfactual | (3) Actual | (4) Difference (2) – (3) | (5) Difference Percent |
|---|---|---|---|---|
| 2020(1) | 27.9162 | 27.9162 | .0009995 | 0.0 |
| 2020(2) | 27.9171 | 27.8299 | .0893255 | 8.9 |
| 2020(3) | 27.9191 | 27.9010 | .0211448 | 2.1 |
| 2020(4) | 27.9221 | 27.9010 | .0241403 | 2.4 |
| 2021(1) | 27.9251 | 27.9162 | .0129805 | 1.2 |
| 2021(2) | 27.9291 | 27.9311 | .0020469 | 0.2 |
| 2021(3) | 27.9331 | 27.9458 | -.0086672 | -0.8 |
| 2021(4) | 27.9371 | 27.9675 | -.0263366 | -2.6 |
| 2022(1) | 27.9411 | 27.9675 | -.0213491 | -2.1 |
| 2022(2) | 27.9461 | 27.9675 | -.0163616 | -1.6 |
| 2022(3) | 27.9511 | 27.9675 | -.0113746 | -1.1 |

In the same way as before, the differences are so-called differences-from-counterfactuals, implying that a positive number is an estimated negative difference between the forecasted counterfactual and actual GDP-numbers, and opposite for positive numbers. The estimated differences for the aggregate models and final form equations, as well as for the comparable studies and estimations, are listed in table 5.6 below.

**Table 5.6:** Difference between counterfactual (no Covid-19) and actual in percent of counterfactual

| | 2020 | 2021 | 2022[2] |
|---|---|---|---|
| Aggregate model, Mainland-Norway | 6.3 | 5.5 | 5.1 |
| Aggregate model, Sweden | 3.6 | 0.3 | -0.4 |
| | | | |
| Final form equation, Mainland-Norway | 5.3 | 3.4 | 1.9 |
| Final form equation, Sweden | 4.5 | 2.1 | 2.0 |
| | | | |
| Corona commission, Mainland-Norway | | | |
| Bjertnæs et al. (2021) | 4.7 | 3.8 | 2.2 |
| von Brasch et al. (2022) | 4.6 | 2.4 | 2.1 |
| | | | |
| Konjunkturinstitutet, Sweden | 3.3 | -0.5 | -1.8 |
| Forecast December 2019 | | | |

Looking at the estimated values of the aggregate models, there is a noteworthy difference between Mainland-Norway and Sweden. The effect is larger in magnitude in Norway for the entirety of the years 2020-2022. Moreover, towards the end of 2021 in Sweden, actual GDP numbers has surpassed the forecasts obtained from the aggregate model. This implies that all the loss from the pandemic has been made up for inside of only three years. In Norway, the situation looks more dire: In 2022, the estimated difference is still 5.1%, indicating that the pandemic and the policy response has resulted in leaving the economy sub performing by over 5% three years after the shock, compared to the estimated forecasts of the aggregate model pre pandemic. The

[2] Until and including the third quarter of 2022

results overall indicate that the more lenient policy regime in Sweden has been beneficial for the economy through the pandemic period.

Compared to the results in the reports on behalf of the corona commission, the aggregate model for GDP in Mainland-Norway estimated a larger difference-from-counterfactual. The combined estimated difference for the years 2020-2022 in von Brasch et al. (2022) was 7.7% lower compared to the findings in my estimations. The estimates from the final form equation were closer to those in the report on the behalf of the corona commission. While the estimated effect in the mentioned three-year period still was stronger than in von Brasch et al. (2022), the combined difference was only 1.5%. However, both results were larger than in the comparable study, implying a potential underestimation in the estimations that has been performed on behalf of the corona commission.

For Sweden, there was a small difference between the aggregate model estimations and the forecast from Konjunkturinstitutet. The aggregate model estimated a larger impact on the economy compared to the forecast, the combined difference for 2020-2022 being 2.5%. Comparing Konjunkturinstitutet's forecast with the final form equation for GDP in Sweden, however, the combined difference was 7.6%. Overall, the results indicate that, similar to the aggregate model for Mainland-Norway GDP, the effect on GDP could be underestimated in the comparable forecasts.

## 5.4 Discussion of the results

In the results presented above, both aggregate models estimated larger difference-from-counterfactuals relative to comparable studies and estimations. This could mean one (or a combination) of two things: The aggregate models constructed in this thesis are overestimating the difference-from counterfactual, or the comparable studies and estimations are underestimating the same difference. It is difficult to argue that the estimates from the aggregate models I have estimated for the purpose of this thesis, are more credible than those made in the reports on behalf of the Norwegian corona commission and by Konjunkturinstitutet. However, through the documentation of my approach presented in the previous chapter, the results from the aggregate models presented are still credible. The differences could as such be due to a combination of difference in sample and methodology, and still be a useful addition to the existing literature.

In a similar study to this, Nymoen (2023) used the Norwegian Aggregate Model (Bårdsen & Nymoen, 2022), to estimate the economic effect of Covid-19 on the Norwegian Mainland economy. The estimated difference between actuals and counterfactuals for 2020 and 2021 are 5.1% and 4.4%, respectively. The comparable values of this paper are 6.3% and 5.5%. This could indicate that there is a tendency towards overestimation in the forecasts that are based on the aggregate models constructed in this paper, but also that the estimations from Bjertnæs et al. (2021) and von Brasch et al. (2022) could be underestimating the loss in GDP for Mainland-Norway.

Because the estimated counterfactuals are dynamic forecasts in both estimation methods, the interpretation and validation are the same for both. The economy should be characterized as accurately as possible by the model specification in the estimation sample and the forecast period, while being invariant to the shock of the pandemic and the corresponding policy response. This rather strong requirement is unlikely to be held for all the equations in the macroeconometric models, and could be one source of difference in the estimations. However, even though the concept of invariance might not hold completely, the property can still hold partly (Nymoen, 2023, p. 18). As such, while the model invariance might not be fully satisfied, the estimations and results are still likely to partly satisfy the requirement.

An important distinction about recovering after economic shocks leading to declines in GDP, is that the loss in GDP is not necessarily fully recovered when it reaches the GDP-level from before the shock. Usually, there is expected growth in GDP that is lost as well. Blytt et al. (2022, p. 16) remarks that both Mainland-Norway and Sweden surpassed the pre-pandemic GDP-level during 2021. This does not, however, account for the added loss of expected growth, meaning that there are still two years' worth of GDP-growth missing as a consequence of the pandemic.

### 5.4.1 Looking beyond the economic effects

The aggregate multiple-equation models constructed for the purpose of this thesis, are an efficient way to analyze the effect of an exogenous shock such as Covid-19 and the corresponding policy responses. The economic impact is measured in loss of GDP, leaving a discussion of this choice necessary. Primarily, loss in GDP does not account for the health costs. Loss of life, reduction in life quality and educational outcome due to social distancing and lockdowns are some examples of non-economic impacts (at least in terms of reduced GDP).

Figure 5.3 below, is showing daily new confirmed Covid-19 deaths per million people for Norway and Sweden. The figure is obtained from Our World in Data, showing data from March 2020 until (and including) January 2023. Moreover, the accumulated number of deaths relative to the decline in economic activity (at the beginning of the second quarter of 2021) for European countries, can be found in figure 5.4.

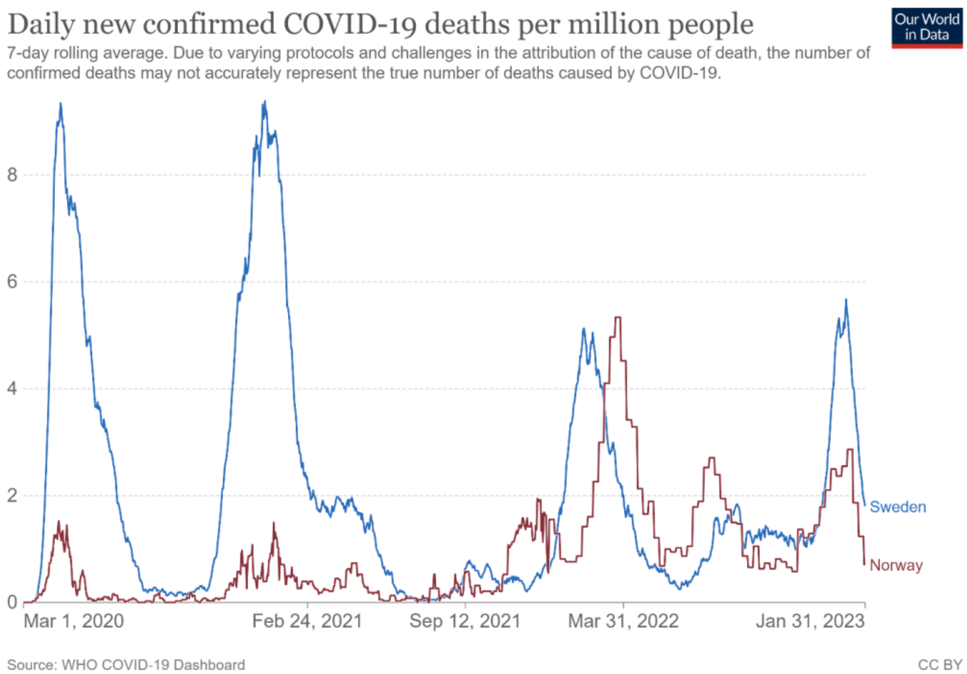


**Figure 5.3:** Daily new confirmed COVID-19 deaths per million people in Norway and Sweden. Source: Mathieu et al. (2023)

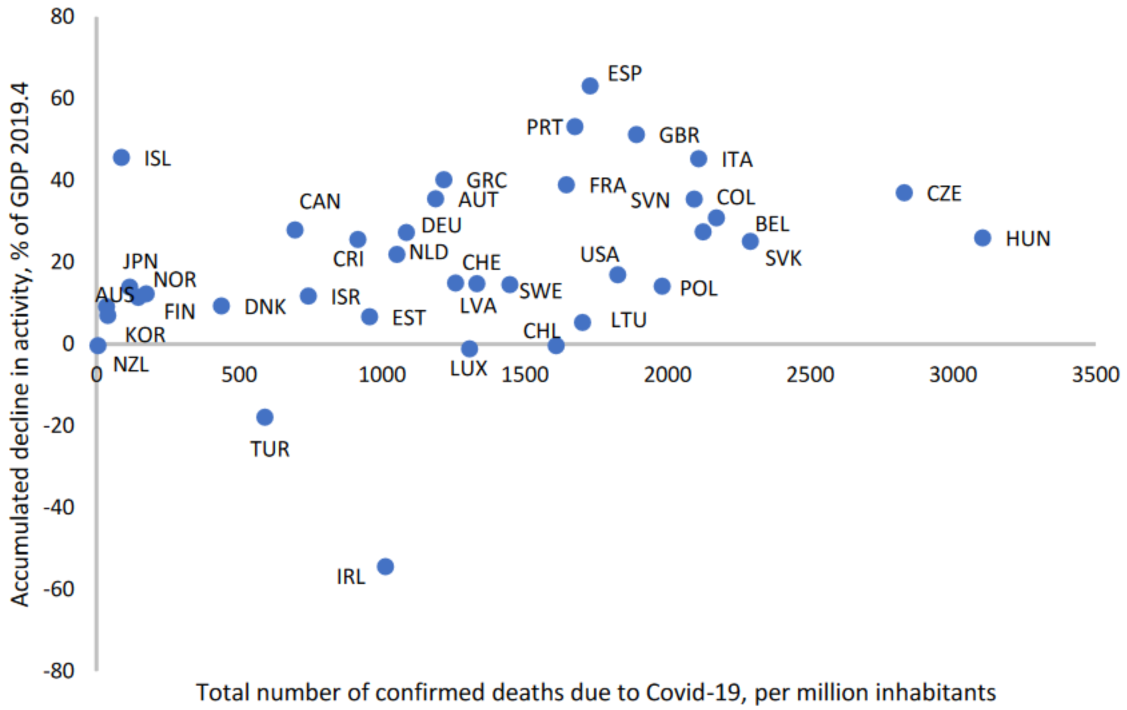


**Figure 5.4:** Accumulated number of Covid-19 deaths and decline in economic activity per 2021(2)[3]. Source: Andersen et al. (2022, p. 8)

The results from the figures are quite clear: The health impact was larger in Sweden. Additionally, the latter figure shows how the difference in policy response has placed Norway and Sweden on rather different places in the figure. Up until the second quarter of 2021, the economic impact was similar, while there was a substantial difference in accumulated deaths (per million inhabitants). Looking back at figures 5.1 and 5.2, Sweden's economic growth were stronger in the quarter succeeding this one, marking another important discovery: Norway's more strict policy response has dampened the long-term economic development compared to Sweden's.

The discussion of the economic worth of a human life is a difficult one, and outside the scope of the research question of this thesis. Moreover, a comparison which would have included all

[3] Accumulated decline in activity is the sum of the difference in activity relative to 2019(4) in the period from 2020(1) to 2021(2), based on seasonally adjusted GDP statistics from www.oecd.library.org.
Mortality data is total deaths due to Covid-19 from start of 2020 to end of March 2021 based on data from www.ourworldindata.org.

possible effects and externalities would be extremely complex and impossible to execute. However, the difference in death tolls is a crucial factor to keep in mind when evaluating the total effect of the pandemic. It also does not change this thesis' role in the literature, which is to contribute by looking at the economic effects of the pandemic. The results can of course be used when discussing the broader picture in more detail.

# Chapter 6: Conclusion

This thesis has investigated the economic effects of Covid-19 and the policy responses to the pandemic in Norway and Sweden. The research question has been empirically analyzed using aggregate macroeconometric models and final form equations to obtain counterfactual outcomes. The aggregate model for Mainland-Norway GDP found a statistically significant long covid effect, while the equivalent for Sweden found no such effect. The final form equation for GDP in Mainland-Norway found a weaker long-term effect compared to the aggregate model. For GDP in Sweden, the final form equation found a larger economic loss compared to what the aggregate model yielded.

Relative to comparable studies, I found a larger economic effect from Covid-19 and the policy responses in both Mainland-Norway and Sweden. For Mainland-Norway, the aggregate model estimations implied an underestimation in the reports written on behalf of the corona commission, while the counterfactual given by the final form equation was more aligned with the latter. For Sweden, both methods estimated a stronger economic impact compared to Konjunkturinstitutet's forecast, but neither yielding a long-term significant effect from the pandemic.

The results overall shows that the strict policy regime in Norway resulted in a significant long-term economic effect, while the more lenient strategy chosen in Sweden resulted in no long-term effects. Considering the results, an important remark is that the purpose of this thesis was to find only the economic effect. Other considerations, such as health effects and death tolls, are not accounted for. The thesis contributes to the literature on economic effects of Covid-19, especially for Sweden, which has no studies using the methods presented here. Moreover, the results may be a useful contribution when doing any further assessments of the effects of the pandemic and policy responses.

# References


Andersen, T. M., Holden, S., & Honkapohja, S. (2022). *Economic consequences of the pandemic—The Nordic countries*.

Bårdsen, G., den Reijer, A., Jonasson, P., & Nymoen, R. (2011). *MOSES: Model of Swedish Economic Studies*.

Bårdsen, G., & Nymoen, R. (2009). *Macroeconometric modelling for policy*.

Bårdsen, G., & Nymoen, R. (2022, September). *NAM – Norwegian Aggregate Model – NORMETRICS*. https://normetrics.no/nam/

Bjertnæs, G. H. M., von Brasch, T., Cappelen, Å., Holden, S., Holmøy, E., Slettebø, O., Sletten, P., & Zhulanova, J. (2021). *COVID-19, tapt verdiskaping og finanspolitikkens rolle. Utredning for Koronakommisjonen*.

Blytt, J. P., Bougroug, A., & Sletten, P. (2022). Økonomisk utvikling gjennom Covid-19. *Statistisk Sentralbyrå, Statistics Norway*.

Castle, L., J., Doornik, J., & Hendry, D. (2012). Model selection when there are multiple breaks. *Journal of Econometrics*, *169*, 239–247.

Doornik, J. (2009). Autometrics. *Castle, J. Shephard, N (Eds.)*, 88–121.

Doornik, J., & Hendry, D. (2018). *Empirical Econometric Modelling PcGive 15. Volume 1*. Timberlake Consultants, London.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987–1007. https://doi.org/10.2307/1912773

Epprecht, C., Guegan, D., & Veiga, A. (2013). *Comparing variable selection techniques for linear regression: LASSO and Autometrics*.

Ericsson, N. R., & MacKinnon, J. G. (2002). Distributions of error correction tests for cointegration. *The Econometrics Journal*, *5*(2), 285–318.

Harvey, A. C. (1981). *Time Series Models*. Philip Allan.

Hendry, D. F., & Doornik, J. A. (2014). *Empirical Model Discovery and Theory Evaluation: Automatic Selection Methods in Econometrics*. The MIT Press. https://www.jstor.org/stable/j.ctt9qf9km

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, *6*(3), 255–259. https://doi.org/10.1016/0165-1765(80)90024-5

Konjunkturinstitutet. (2019). Konjunkturläget december 2019. *Konjunkturinstitutet*, *December 2019*, 84.

Mackinnon, J. (1990). Critical Values for Cointegration Tests. *Long-Run Economic Relationships*.

Macrobond. (2023). *Products—Economic Data | Macrobond Financial*. https://www.macrobond.com/products

Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Giattino, C., Hasell, J., Macdonald, B., Dattani, S., Beltekian, D., Ortiz-Ospina, E., & Roser, M. (2023). Coronavirus Pandemic (COVID-19). *Our World in Data*. https://ourworldindata.org/covid-deaths

Muhammadullah, S., Urooj, A., Khan, F., Alshahrani, M. N., Alqawba, M., & Al-Marzouki, S. (2022). Comparison of Weighted Lag Adaptive LASSO with Autometrics for Covariate Selection and Forecasting Using Time-Series Data. *Complexity*, *2022*, e2649205. https://doi.org/10.1155/2022/2649205

*National accounts—Statistics Norway*. (2023). SSB. https://www.ssb.no/en/nasjonalregnskap-og-konjunkturer/nasjonalregnskap/statistikk/nasjonalregnskap

Nymoen, R. (2019). *Dynamic econometrics for empirical macroeconomic modelling*. World Scientific.

Nymoen, R. (2023). Economic Covid-19 effects analysed by macro econometric models—The case of Norway. *National Accounting Review*, *5*(1), 1–22. https://doi.org/10.3934/NAR.2023001

Pesaran, M., & Smith, R. (2016). Counterfactual analysis in macroeconometrics: An empirical investigation into the effects of quantitative easing. *Research in Economics*, *70*(2), 262–280.

R Core Team. (2022). *R: The R Project for Statistical Computing* (4.2.1). R Foundation for Statistical Computing. https://www.r-project.org/

Regeringskansliet. (2020). *Ministry of Finance economic forecasts*. Regeringskansliet; Regeringen och Regeringskansliet. https://www.government.se/government-of-sweden/ministry-of-finance/central-government-budget/ministry-of-finance-economic-forecasts/

Riksbanken. (2023). *Effective exchange rate index—KIX and TCW*. https://www.riksbank.se/en-gb/statistics/search-interest--exchange-rates/explanation-of-the-series/effective-exchange-rate-indices/

Stock, J. H., & Watson, M. W. (2015). *Introduction to econometrics* (updated 3. ed., global ed). Pearson.

von Brasch, T., Cappelen, Å., Holden, S., Lindstrøm, E. L., & Skretting, J. (2022). COVID-19, tapt verdiskaping og finanspolitikkens rolle. *RAPPORTER/REPORTS 2022/15*.

# Appendices

## Appendix 1: Supplementary topics in time series econometrics

### A.1.1 Difference equations

Dynamic econometric modelling is built upon the mathematical theory of difference equations (Nymoen, 2019, p. 103). Difference equations can be deterministic or stochastic. Deterministic difference equations are the discrete time counterpart to differential equations, which is used in economic theory for continuous growth, such as the Solow model. A deterministic difference equation may be either homogenous or inhomogeneous.

A linear homogenous difference equation can be written as:

$$a_0 x_t + a_1 x_{t-1} + \cdots + a_p x_{t-p} = 0 \qquad (A.1)$$

While a linear inhomogeneous difference equation can be written as:

$$a_0 x_t + a_1 x_{t-1} + \cdots + a_p x_{t-p} = b_t \qquad (A.2)$$

Below, $C_j$ is denoting arbitrary constants and $\lambda_j$ denotes the distinct roots of the characteristic equation $p(\lambda) = 0$. In the case of p distinct roots, the solution of the inhomogeneous equation is:

$$x_t = C_1 \lambda_1^t + C_2 \lambda_2^t + \cdots + C_p \lambda_p^t + x_t^s \qquad (A.3)$$

$$x_t = x_t^h + x_t^s, \qquad (A.4)$$

$$where\ x_t^h\ is\ the\ general\ solution\ of\ (I)\ and\ x_t^s\ is\ a\ special\ solution\ of\ (A.2).$$

A stochastic difference equation of order p with constants coefficients is denoted as:

$$Y_t = \phi_0 + +\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \qquad (A.5)$$

Where $Y_{t-i}, i = 0,1, \dots, p$ and $\varepsilon_t$ are random variables and $\phi_i$ are constants. The error terms are assumed to be independent and identically distributed with expectation zero and variance $\sigma_\varepsilon^2$, which can be annotated as:

$$\varepsilon_t \sim IID(0, \sigma_\varepsilon^t)\ \forall t \qquad (A.6)$$

The stochastic equation (A.5) and the deterministic equation (A.2) has the same mathematical properties. Consequently, the solution can be obtained in the same way as with the inhomogeneous equation. The homogenous part of the stochastic difference equation is associated with the characteristic polynomial equation defined as:

$$p(y) = y^p - \phi_1 y^{p-1} - \cdots - \phi_p \qquad (A.7)$$

Where $\lambda$ can be defined as a characteristic root in the characteristic equation $p(\lambda) = 0$:

$$\lambda^p - \phi_1 \lambda^{p-1} - \cdots - \phi_p = 0 \qquad (A.8)$$

As with the general solution for the deterministic equation, the general solution $Y_t$ can be denoted as:

$$Y_t = Y_t^h + Y_t^s, \qquad (A.9)$$

$$\text{where } Y_t^s \text{ denotes a special solution of } (A.5).$$

### A.1.2 Identification; order and rank condition

The ability to estimate the causal effect of one variable on another is referred to as identification. If a relationship is identified, it is possible to estimate the effect of a change in one variable on another. If, however, a relationship is not identified, it is not possible to estimate the causal effect of the relationship.

The first important concept of identification of a system is the order condition. It states that in a system of linear equations, the number of variables must be at least as large as the number of equations. A system with more equations than variables is called overidentified, while a system with less equations than variables is said to be underidentified.

A second important concept of identification is the rank condition. The rank condition states that in a system of linear equations, the number of linearly independent equations must be at least as large as the number of unknown parameters. If the rank condition is not satisfied, the parameters cannot be identified.

The order and conditions together determine whether a system of linear equations is identified. If both conditions are satisfied, the system is fully identified and the parameters in the system can be estimated. The order and rank conditions are important criteria for determining the identification of a system of linear equations. In the software used in this thesis, namely OxMetrics 8, a system will not be estimated without these conditions satisfied.

### A.1.3 Exogeneity

There are three different degrees of exogeneity; weak, strong, and super. A variable is defined as weakly exogenous if it can be estimated consistently and efficiently by only considering the conditional model, disregarding the rest of the system (Nymoen, 2019, p. 303). It follows from this that the residuals in the conditional and marginal are uncorrelated. A necessary condition for estimation is weak exogeneity, and as such, all variables in the framework of this thesis is assumed to be weakly exogenous.

A variable is strongly exogenous if it is both weakly exogenous and Granger non-causal. The latter means that the variable is not Granger-caused by lags of the endogenous variable. This (one-way causality) is a necessary condition for forecasting: The cause must come before the effect.

Finally, a variable is super exogenous if it is both weakly exogenous and the corresponding coefficient is invariant to structural breaks. This implies that any significant dummies for structural breaks in the marginal model are insignificant in the conditional model.

## Appendix 2: Table with critical values used in EG-tests and ADF-tests

**Table A.1:** Critical Values used in EG-tests and ADF-tests. Source: MacKinnon (1990, p. 9)

| $N$ | Variant | Level | Obs. | $\beta_\infty$ | (s.e.) | $\beta_1$ | $\beta_2$ |
|---|---|---|---|---|---|---|---|
| 1 | no constant | 1% | 600 | −2.5658 | (0.0023) | −1.960 | −10.04 |
| | | 5% | 600 | −1.9393 | (0.0008) | −0.398 | |
| | | 10% | 560 | −1.6156 | (0.0007) | −0.181 | |
| 1 | no trend | 1% | 600 | −3.4336 | (0.0024) | −5.999 | −29.25 |
| | | 5% | 600 | −2.8621 | (0.0011) | −2.738 | −8.36 |
| | | 10% | 600 | −2.5671 | (0.0009) | −1.438 | −4.48 |
| 1 | with trend | 1% | 600 | −3.9638 | (0.0019) | −8.353 | −47.44 |
| | | 5% | 600 | −3.4126 | (0.0012) | −4.039 | −17.83 |
| | | 10% | 600 | −3.1279 | (0.0009) | −2.418 | −7.58 |
| 2 | no trend | 1% | 600 | −3.9001 | (0.0022) | −10.534 | −30.03 |
| | | 5% | 600 | −3.3377 | (0.0012) | −5.967 | −8.98 |
| | | 10% | 600 | −3.0462 | (0.0009) | −4.069 | −5.73 |
| 2 | with trend | 1% | 600 | −4.3266 | (0.0022) | −15.531 | −34.03 |
| | | 5% | 560 | −3.7809 | (0.0013) | −9.421 | −15.06 |
| | | 10% | 600 | −3.4959 | (0.0009) | −7.203 | −4.01 |
| 3 | no trend | 1% | 560 | −4.2981 | (0.0023) | −13.790 | −46.37 |
| | | 5% | 560 | −3.7429 | (0.0012) | −8.352 | −13.41 |
| | | 10% | 600 | −3.4518 | (0.0010) | −6.241 | −2.79 |
| 3 | with trend | 1% | 600 | −4.6676 | (0.0022) | −18.492 | −49.35 |
| | | 5% | 600 | −4.1193 | (0.0011) | −12.024 | −13.13 |
| | | 10% | 600 | −3.8344 | (0.0009) | −9.188 | −4.85 |
| 4 | no trend | 1% | 560 | −4.6493 | (0.0023) | −17.188 | −59.20 |
| | | 5% | 560 | −4.1000 | (0.0012) | −10.745 | −21.57 |
| | | 10% | 600 | −3.8110 | (0.0009) | −8.317 | −5.19 |
| 4 | with trend | 1% | 600 | −4.9695 | (0.0021) | −22.504 | −50.22 |
| | | 5% | 560 | −4.4294 | (0.0012) | −14.501 | −19.54 |
| | | 10% | 560 | −4.1474 | (0.0010) | −11.165 | −9.88 |
| 5 | no trend | 1% | 520 | −4.9587 | (0.0026) | −22.140 | −37.29 |
| | | 5% | 560 | −4.4185 | (0.0013) | −13.641 | −21.16 |
| | | 10% | 600 | −4.1327 | (0.0009) | −10.638 | −5.48 |
| 5 | with trend | 1% | 600 | −5.2497 | (0.0024) | −26.606 | −49.56 |
| | | 5% | 600 | −4.7154 | (0.0013) | −17.432 | −16.50 |
| | | 10% | 600 | −4.4345 | (0.0010) | −13.654 | −5.77 |
| 6 | no trend | 1% | 480 | −5.2400 | (0.0029) | −26.278 | −41.65 |
| | | 5% | 480 | −4.7048 | (0.0018) | −17.120 | −11.17 |
| | | 10% | 480 | −4.4242 | (0.0010) | −13.347 | |
| 6 | with trend | 1% | 480 | −5.5127 | (0.0033) | −30.735 | −52.50 |
| | | 5% | 480 | −4.9767 | (0.0017) | −20.883 | −9.05 |
| | | 10% | 480 | −4.6999 | (0.0011) | −16.445 | |

## Appendix 3: Scripts for aggregate models and final form equations

The datasets and script files used in this thesis are available from the author upon request. A variable list with complementary variable names from the raw data is also added below.

### A.3.1 Complementary variable list

**Norway**

Endogenous:

**Y** = GDP Norway, market values, fixed prices, mill. NOK.

**YF** = GDP Mainland-Norway, market values, fixed prices, mill. NOK.

**G** = Public consumption expenditure, fixed prices, mill. NOK (NAM: CO) + Gross fixed capital formation, general government, fixed prices, mill. NOK (NAM: JO).

**EMI** = Export market indicator, index.

**JOIL** = Gross fixed capital formation (GFCF), oil and gas production and pipeline transportation (NAM: JOIL1), and related services (NAM: JOIL2), fixed prices, mill. NOK.

Exogenous:

**P** = Consumer Price Index (CPI).

**INF** = CPI inflation rate, log approximated ($INF = 100 * p_t - p_{t-4}$, where lowercase p denotes the natural logarithm of CPI)

**RL** = Average interest rate on total bank loans, percent. Equal to $\mathbf{R_L}$ in re-estimation of GDP-equation in NAM-2009.

**RLINF**= Real interest rate, difference between average interest rate on bank loans and log approximated CPI inflation, percent.

**L** = Domestic credit to general public, K2-indicator. (NAM: K2)

**ARBDAG** = Number of working days per quarter.

**V** = Trade-weighted nominal value of the krone based on import-shares of trading countries. (NAM: CPIVAL)

**P*** = Foreign consumer price index (trade weighted). (NAM: PCKONK)

**π** = CPI inflation ($\pi = \frac{P_t - P_{t-4}}{P_{t-4}} * 100$)

**Sweden**

Endogenous:

**Y** = Gross Domestic Product, Total, Constant Prices, Seasonally Adjusted (SA), Market Prices, SEK.

**G** = Central Government Budget, Expenditures, Total, SEK.

**EMI** = Foreign GDP, Closest trading partners (China, Denmark, Euro Area, Norway, Poland, UK, USA), KIX-weighted. Equal to **YF** in re-estimation of GDP-equation in MOSES.

Exogenous:

**CPI** = Consumer Price Index (Riksbank Classification), Total, Index.

**RL** = Deposits & Loans, Banks, Lending Rates, By Entity, to Households including NPISH, Period Ending Stock, All Accounts, percent.

**INF** = CPI inflation rate ($INF = 100 * \frac{\Delta_4 CPI_t}{CPI_{t-4}}$). Equal to **π** in re-estimation of GDP-equation in MOSES.

**RLINF** = Real CPI interest rate, percent.

**ARBDAG** = Number of Working Days, Per Quarter.

### A.3.2 Script for re-estimation of GDP-equation in NAM-2009

```
// Re-estimation of GDP-equation in NAM-2009 //

// Variable definitions:
// Y = Total GDP in market values, fixed prices, mill. NOK
// P = Consumer Price Index (CPI)
// P* = Foreign consumer price index, trade weighted (PCKONK)
// INF = CPI inflation rate
// RL = Average interest rate on total bank loans, percent
// RLINF = Real interest rate, difference between average interest rate on bank loans and CPI inflation, percent.
// L = Domestic credit to general public, K2-indicator (K2)
// V = Trade weighted nominal value of NOK based on import shares of trading countries (CPIVAL)
// G = Public consumption expenditure, fixed prices, mill. NOK (CO) + Gross fixed capital formation, general government, fixed prices, mill. NOK (JO)

// Loading dataset
module("PcGive");
package("PcGive", "Single-equation");
loaddata("NAMdatabase_desember2022.xlsx");
usedata("NAMdatabase_desember2022.xlsx");

// Constructing required variables
algebra{
V = CPIVAL;
INF = ((CPI - lag(CPI,4)) / lag(CPI,4))*100; // π
RLINF = RL - INF; // RL - π
"l-p" =  log(K2) - log(CPI);
LY = log(Y);
DLY = dlog(Y);
LG = log(CO+JO); // Combining public expenditure and consumption
DLG = diff(LG,1);
LV = log(V);
"LP*" = log(PCKONK);
LP = log(CPI);
"Dl-p" = diff("l-p",1);
"LV+LP*-LP" = LV + LP* - LP;
}

// 17.44 on unrestricted form
module("PcGive");
package("PcGive", "Single-equation");
usedata("NAMdatabase_desember2022.xlsx");
system
{
    Y = DLY;
    Z = Constant, DLY_1, LG_1, "LV+LP*-LP_1", RLINF_1, LY_2, DLG,
        "Dl-p_1";
}
estimate("OLS", 1988, 2, 2007, 1);

// Re-estimation of 17.44 in NAM-2009
// Change: Using constant instead of estimated mean of the ECM-part

// Creating the ECM-part
algebra{
ECM = lag(LY,2) - 0.9*lag(LG,1) - 0.16*lag("LV+LP*-LP",1) + 0.006*lag(RLINF,1);
}

// 17.44 on ECM form
module("PcGive");
```

```
package("PcGive", "Single-equation");
usedata("NAMdatabase_desember2022.xlsx");
system
{
    Y = DLY;
    Z = Constant, ECM, DLY_1, DLG, "Dl-p_1";
}
estimate("OLS", 1988, 2, 2007, 1);
```

### A.3.3 Script for re-estimation of GDP-equation in MOSES

```
// Re-estimation of GDP-equation in MOSES //

// Variable definitions
// Y = Gross Domestic Product, Total, Constant Prices, SA, Market Prices, SEK
// G = Central Government Budget, Expenditures, Total, SEK
// GOV = Goverment expenditure term, (G/Y)_1*log(G), SEK
// CPI = Consumer Price Index (Riksbank Classification), Total, Index
// INF = CPI inflation rate
// RL = Deposits & Loans, Banks, Lending Rates, By Entity, to Households including NPISH, Period Ending Stock, All Accounts
// RLINF = Real interest rate, difference between average interest rate on bank loans and CPI inflation, percent
// YF = Foreign GDP, Closest trading partners (China, Denmark, Euro Area, Norway, Poland, UK, USA), KIX-weighted

// Loading dataset
module("PcGive");
package("PcGive", "Single-equation");
loaddata("GDP_Sweden_v6.csv");
usedata("GDP_Sweden_v6.csv");

// Constructing required variables
algebra{
LY = log(Y);
DLY = dlog(Y);
LYF = log(YF);
DLYF = dlog(YF);
"LY-LYF" = LY - LYF;
G = CO;
GY = G/Y;
GY_1 = lag(GY, 1);
LG = log(G);
GOV = GY_1*LG;
DGOV = diff(GOV, 1);
INF = (100*diff(CPI,4))/lag(CPI,4); // π
RLINF = RL - INF; // RL - π
}

// Equation 53 in MOSES WP
module("PcGive");
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v6.csv");
system
{
    Y = DLY;
    Z = "LY-LYF_1", DGOV, DLY_1, RLINF_4, DLYF, DLYF_2, Constant;
}
estimate("OLS", 1997, 01, 2009, 4); // Sample size used in MOSES
// Shift dummy x_st in original GDP-equation "is proxying a shift in export shares from 2004(1).

// Eq. 53 in MOSES WP, but using larger sample size
module("PcGive");
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v6.csv");
system
{
    Y = DLY;
    Z = "LY-LYF_1", DGOV, DLY_1, RLINF_4, DLYF, DLYF_2, Constant;
}
estimate("OLS", 1996, 01, 2020, 4);
```

Page: 1 of 1

### A.3.4 Script for modelling of GDP in Mainland-Norway

```
// Aggregate model for Norway using NAM data - GDP-equation //

// Variable definitions:
// Y = Total GDP in market values, fixed prices, mill. NOK
// YF = GDP mainland Norway, market values, fixed prices, mill. NOK
// P = Consumer Price Index (CPI)
// INF = Log approximation of CPI inflation
// RL = Average interest rate on total bank loans, percent
// RLINF = Real interest rate, difference between average interest rate on bank loans and log approximated CPI inflation, percent.
// L = Domestic credit to general public, K2-indicator (K2)
// G = Public consumption expenditure, fixed prices, mill. NOK (CO) + Gross fixed capital formation, general government, fixed prices, mill. NOK (JO)
// EMI = Export market indicator, index
// JOIL = Gross fixed capital formation (GFCF), oil and gas production and pipeline tranportation (JOIL1), and related services (JOIL2), fixed prices. JOIL = JOIL1 + JOIL2
// ARBDAG = Number of working days per quarter

// Loading dataset
module("PcGive");
package("PcGive", "Single-equation");
loaddata("NAMdatabase_desember2022.xlsx");
usedata("NAMdatabase_desember2022.xlsx");


// Constructing required variables
algebra{
P = CPI;
LP = log(P);
INF = 100*(LP)-lag(LP,4); // Log approximation of inflation
RLINF = RL - INF;
RLINF = RLINF/100; // Transforms the variable into the rate of real inflation
LY = log(Y); // Total GDP
DLY = dlog(Y);
LYF = log(YF); // Mainland-GDP
DLYF = dlog(YF);
LG = log(CO+JO);
DLG = diff(LG,1);
"l-p" =  log(K2) - log(CPI);
"Dl-p" = diff("l-p",1);
LEMI =log(EMI);    // Foreign market indicator
DLEMI =dlog(EMI);
LARBDAG=log(ARBDAG);
DLARBDAG=dlog(ARBDAG);
JOIL=JOIL1+JOIL2;
DLJOIL= dlog(JOIL);
}

// First final model equation for GDP - with trend variables in the GUM
system
{
    Y = DLYF;
    Z = Constant,
        DLYF_1,DLYF_2,DLYF_3, DLYF_4,
        LYF_1, LG_1, LEMI_1,
        DLARBDAG, LARBDAG_1,
        DLG,DLG_1,DLG_2, DLG_3,DLG_4,
        DLJOIL,DLJOIL_1,DLJOIL_2, DLJOIL_3, DLJOIL_4,
        DLEMI,DLEMI_1,DLEMI_2,DLEMI_3,DLEMI_4,
        RLINF, RLINF_1, RLINF_2,
```

```
        "Dl-p_1",
        CSeasonal,CSeasonal_1,CSeasonal_2;
}

// Applying Autometrics
autometrics(0.01, "IIS", 1);
estimate("OLS", 1988, 2, 2019, 4);

// Equation estimated using Autometrics - with trend variables in the GUM
module("PcGive");
package("PcGive", "Single-equation");
usedata("NAMdatabase_desember2022.xlsx");
system
{
    Y = DLYF;
    Z = DLYF_1, DLYF_4, LYF_1, LG_1, LEMI_1, DLARBDAG, LARBDAG_1, DLG,
        DLG_1, DLJOIL, DLJOIL_1, DLJOIL_2, DLEMI, CSeasonal,
        CSeasonal_2, "I:1991(4)", "I:2006(4)";
}
// Formulation of the GUM (commented out)
/*
system
{
    "Y" = "DLYF", "DLYF_1", "DLYF_2", "DLYF_3", "DLYF_4";
    "Z" = "Constant", "LYF_1", "LG_1", "LEMI_1", "DLARBDAG", "LARBDAG_1", "DLG", "DLG_1",
        "DLG_2", "DLG_3", "DLG_4", "DLJOIL", "DLJOIL_1", "DLJOIL_2", "DLJOIL_3",
        "DLJOIL_4", "DLEMI", "DLEMI_1", "DLEMI_2", "DLEMI_3", "DLEMI_4", "RLINF",
        "RLINF_1", "RLINF_2", "Dl-p_1", "CSeasonal", "CSeasonal_1", "CSeasonal_2";
}
autometrics(0.01, "IIS", 1);
*/
estimate("OLS", 1988, 2, 2019, 4);

// Creating the ECM-equation for the Long-run relationship between yf, g and emi
algebra{
LYF_ECM = LYF - LG*(0.084/0.176) - LEMI*(0.056/0.176);
LYF_ECM_1 = lag(LYF_ECM, 1);
}

// Second final model equation for GDP - with ECM-variable
system
{
    Y = DLYF;
    Z = Constant,
        DLYF_1,DLYF_2,DLYF_3, DLYF_4,
        LYF_ECM_1,
        DLARBDAG, LARBDAG_1,
        DLG,DLG_1,DLG_2,DLG_3,DLG_4,
        DLJOIL,DLJOIL_1,DLJOIL_2, DLJOIL_3, DLJOIL_4,
        DLEMI,DLEMI_1,DLEMI_2,DLEMI_3,DLEMI_4,
        RLINF, RLINF_1, RLINF_2,
        "Dl-p_1",
        CSeasonal,CSeasonal_1,CSeasonal_2;
}

// Applying Autometrics
autometrics(0.01, "IIS", 1);
estimate("OLS", 1988, 2, 2019, 4);

// Equation estimated using Autometrics - with ECM-variable and no trend variables
```

```
module("PcGive");
package("PcGive", "Single-equation");
usedata("NAMdatabase_desember2022.xlsx");
system
{
    Y = DLYF;
    Z = DLYF_1, DLYF_4, LYF_ECM_1, DLARBDAG, LARBDAG_1, DLG, DLJOIL,
        DLJOIL_1, DLJOIL_2, DLEMI, CSeasonal_2, "I:1996(4)", "I:1998(2)",
        "I:2001(3)", "I:2006(4)", "I:2007(4)", "I:2011(1)", "I:2012(2)";
}
// Formulation of the GUM (commented out)
/*
system
{
    "Y" = "DLYF", "DLYF_1", "DLYF_2", "DLYF_3", "DLYF_4";
    "Z" = "Constant", "LYF_ECM_1", "DLARBDAG", "LARBDAG_1", "DLG", "DLG_1", "DLG_2", "DLG_3",
        "DLG_4", "DLJOIL", "DLJOIL_1", "DLJOIL_2", "DLJOIL_3", "DLJOIL_4", "DLEMI",
        "DLEMI_1", "DLEMI_2", "DLEMI_3", "DLEMI_4", "RLINF", "RLINF_1", "RLINF_2",
        "Dl-p_1", "CSeasonal", "CSeasonal_1", "CSeasonal_2";
}
autometrics(0.01, "IIS", 1);
*/
estimate("OLS", 1988, 2, 2019, 4);

// Static regression with LYF, LG and LEMI:
package("PcGive", "Single-equation");
usedata("NAMdatabase_desember2022.xlsx");

system
{
    Y = LYF;
    Z = Constant, LG, LEMI;
}
estimate("OLS", 1988, 2, 2022, 3); // Covid Q's included, as even those quarters shouldn't destroy a long-run relationship if it is there.

// Based on the results, estimating the static ECM-variable
algebra{
LYF_static=LYF -LG*0.45-LEMI*0.30;
}

// Marginal model - DLG
module("PcGive");
package("PcGive", "Single-equation");
usedata("NAMdatabase_desember2022.xlsx");

system
{
    "Y" = "DLG", "DLG_1", "DLG_2", "DLG_3", "DLG_4", "DLG_5", "DLG_6";
    "Z" = "LARBDAG_1", "DLARBDAG", "LARBDAG", "DLARBDAG_1", "CSeasonal", "CSeasonal_1", "CSeasonal_2";
    "U" = "Constant";
}
autometrics(0.01, "IIS", 1);
estimate("OLS", 1988, 2, 2019, 4);

system
{
    Y = DLG;
    Z = DLG_1, DLG_2, DLG_3, DLARBDAG, "I:1991(4)", "I:2007(4)",
```

```
        "I:2008(3)", "I:2009(4)", "I:2010(4)", "I:2011(4)", "I:2012(1)";
    U = Constant;
}
estimate("OLS", 1988, 2, 2019, 4);

// Marginal model - DLEMI
module("PcGive");
package("PcGive", "Single-equation");
usedata("NAMdatabase_desember2022.xlsx");

system
{
    "Y" = "DLEMI", "DLEMI_1", "DLEMI_2", "DLEMI_3", "DLEMI_4", "DLEMI_5", "DLEMI_6";
    "Z" = "LARBDAG", "LARBDAG_1", "DLARBDAG", "DLARBDAG_1", "CSeasonal", "CSeasonal_1", "CSeasonal_2";
    "U" = "Constant";
}
autometrics(0.01, "IIS", 1);
estimate("OLS", 1988, 2, 2019, 4);

system
{
    Y = DLEMI;
    Z = DLEMI_1, "I:2008(4)", "I:2009(1)", "I:2009(3)";
    U = Constant;
}
estimate("OLS", 1988, 2, 2019, 4);

// Marginal model - DLJOIL
module("PcGive");
package("PcGive", "Single-equation");
usedata("NAMdatabase_desember2022.xlsx");

system
{
    "Y" = "DLJOIL", "DLJOIL_1", "DLJOIL_2", "DLJOIL_3", "DLJOIL_4", "DLJOIL_5", "DLJOIL_6";
    "Z" = "LARBDAG", "LARBDAG_1", "DLARBDAG", "DLARBDAG_1", "CSeasonal", "CSeasonal_1", "CSeasonal_2";
    "U" = "Constant";
}
autometrics(0.01, "IIS", 1);
estimate("OLS", 1988, 2, 2019, 4);

system
{
    Y = DLJOIL;
    Z = DLJOIL_1, DLJOIL_4, CSeasonal, "I:1988(4)", "I:1989(1)",
        "I:1990(2)", "I:1991(3)", "I:1994(4)", "I:1996(1)", "I:1996(3)",
        "I:1996(4)", "I:1997(1)", "I:1997(2)", "I:1999(4)", "I:2000(1)";
    U = Constant;
}
estimate("OLS", 1988, 2, 2019, 4);

// Final model system: Conditional + marginal models
module("PcGive");
package("PcGive", "Multiple-equation");
usedata("NAMdatabase_desember2022.xlsx");
system
{
    Y = DLYF, DLG, DLEMI, DLJOIL;
```

```
    Z = Constant, LYF_1,
        DLYF_1, DLYF_4, LYF_ECM_1, DLARBDAG, LARBDAG_1, DLJOIL_1,       // From DLYF-equation
        DLJOIL_2, CSeasonal_2, "I:1996(4)", "I:1998(2)", "I:2001(3)",
        "I:2006(4)", "I:2007(4)", "I:2011(1)", "I:2012(2)",
        DLG_1, DLG_2, DLG_3, "I:1991(4)", "I:2008(3)", "I:2009(4)",     // From DLG-equation
        "I:2010(4)", "I:2011(4)", "I:2012(1)",
        DLEMI_1, "I:2008(4)", "I:2009(1)", "I:2009(3)",                 // From DLEMI-equation
        DLJOIL_4, CSeasonal, "I:1988(4)", "I:1989(1)",
        "I:1990(2)", "I:1991(3)", "I:1994(4)", "I:1996(1)",             // From DLJOIL-equation
        "I:1996(3)", "I:1997(1)", "I:1997(2)", "I:1999(4)", "I:2000(1)";
    I = LYF_ECM, LYF;
}
model
{
    "DLYF" = "DLYF_1", "DLYF_4", "LYF_ECM_1", "DLARBDAG", "LARBDAG_1", "DLG", "DLJOIL", "DLJOIL_1",
        "DLJOIL_2", "DLEMI", "CSeasonal_2", "I:1996(4)", "I:1998(2)", "I:2001(3)", "I:2006(4)",
        "I:2007(4)", "I:2011(1)", "I:2012(2)";
    "DLG" = "Constant", "DLG_1", "DLG_2", "DLG_3", "DLARBDAG", "I:1991(4)", "I:2007(4)", "I:2008(3)",
        "I:2009(4)", "I:2010(4)", "I:2011(4)", "I:2012(1)";
    "DLEMI" = "Constant", "DLEMI_1", "I:2008(4)", "I:2009(1)", "I:2009(3)";
    "DLJOIL" = "Constant", "DLJOIL_1", "DLJOIL_4", "CSeasonal", "I:1988(4)", "I:1989(1)", "I:1990(2)", "I:1991(3)",
        "I:1994(4)", "I:1996(1)", "I:1996(3)", "I:1996(4)", "I:1997(1)", "I:1997(2)", "I:1999(4)",
        "I:2000(1)";
    "LYF_ECM" = "DLYF", "DLG", "DLEMI", "LYF_ECM_1";
    "LYF" = "DLYF", "LYF_1";
}
estimate("FIML", 1988, 2, 2019, 4);

// Forecasting
forecast(11); // 2020(1) - 2022(3) // from Covid shock till end of sample

// Final form equation for DLY
module("PcGive");
package("PcGive", "Single-equation");
usedata("NAMdatabase_desember2022.xlsx");

system
{
    Y = DLYF;
    Z = Constant,
        DLYF_1,DLYF_2,DLYF_3, DLYF_4,
        DLYF_5, DLYF_6, DLYF_7, DLYF_8,
        DLYF_9, DLYF_10, DLYF_11, DLYF_12,
        CSeasonal,CSeasonal_1,CSeasonal_2;
}

// Applying Autometrics
autometrics(0.01, "IIS", 1);
estimate("OLS", 1988, 2, 2019, 4);

// Adding the identity equation of LY to the final form equation
module("PcGive");
package("PcGive", "Multiple-equation");
usedata("NAMdatabase_desember2022.xlsx");

system
```

```
{
    Y = DLYF;
    Z = Constant, LYF_1, DLYF_1, DLYF_12,
        CSeasonal, CSeasonal_1, CSeasonal_2, "I:1997(2)";
    I = LYF;
}
model
{   "DLYF" = "Constant", "DLYF_1", "DLYF_12",
        "CSeasonal", "CSeasonal_1", "CSeasonal_2", "I:1997(2)";
    "LYF" = "DLYF", "LYF_1";
}
estimate("FIML", 1988, 2, 2019, 4);

// Final form equation forecast
forecast(11);
```

### A.3.5 Script for modelling of GDP in Sweden

```
// Aggregate model for Sweden using Macrobond data //

// Variable definitions
// Y = Gross Domestic Product, Total, Constant Prices, SA, Market Prices, SEK
// GO = Central Government Budget, Budget Balance, Total, SEK
// G = Central Government Budget, Expenditures, Total, SEK
// JO = Central Government Budget, Revenues, Total, SEK
// CPI = Consumer Price Index (Riksbank Classification), Total, Index
// RL = Deposits & Loans, Banks, Lending Rates, By Entity, to Households including NPISH, Period Ending Stock, All Accounts
// EMI = Foreign GDP, Closest trading partners (China, Denmark, Euro Area, Norway, Poland, UK, USA), KIX-weighted. Equal to YF in re-estimation of MOSES-2011; renamed to equal the Norwegian model
// INF = CPI inflation rate
// RLINF = Real CPI interest rate
// ARBDAG = Number of Working Days, Per Quarter

// Loading dataset
module("PcGive");
package("PcGive", "Single-equation");
loaddata("GDP_Sweden_v8.csv");
usedata("GDP_Sweden_v8.csv");

algebra{
LY = log(Y);
DLY = dlog(Y);
EMI = YF; // Renaming
LEMI = log(EMI);
DLEMI = dlog(EMI);
G = CO;
LG = log(G);
DLG = dlog(G);
INF = (100*diff(CPI,4))/lag(CPI,4); // π
RLINF = RL - INF; // RL - π
RLINF = RLINF/100; // Transforms the variable into the rate of real inflation
LARBDAG = log(ARBDAG);
DLARBDAG = dlog(ARBDAG);
"LY-LEMI" = LY-LEMI;
}

system
{
    "Y" = "DLY";
    "Z" = "Constant",
        "DLY_1", "DLY_2", "DLY_3", "DLY_4",
        LY_1, LG_1, LEMI_1,
        "DLEMI", "DLEMI_1", "DLEMI_2", "DLEMI_3", "DLEMI_4",
        "DLG", "DLG_1", "DLG_2", "DLG_3", "DLG_4",
        "RLINF", "RLINF_1", "RLINF_2", "RLINF_3", "RLINF_4",
        "LARBDAG", "LARBDAG_1", "DLARBDAG", "DLARBDAG_1",
        "LY-LEMI", "LY-LEMI_1",
        "CSeasonal", "CSeasonal_1", "CSeasonal_2";
}
autometrics(0.01, "IIS", 1);
estimate("OLS", 1996, 3, 2019, 4);

// The equation given by Autometrics - with trend variables
module("PcGive");
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v8.csv");
```

Page: 1 of 5

```
system
{
    Y = DLY;
    Z = Constant, LY_1, LG_1, LEMI_1, DLEMI, DLEMI_1, DLEMI_2, DLG_1,
        DLG_2, DLG_3, "I:2008(4)", "I:2011(3)", "I:2011(4)";
}
estimate("OLS", 1996, 3, 2019, 4);

// Creating the ECM-equation for the long-run relationship between yf, g and emi
algebra{
LY_ECM = LY - LG*(0.047/0.149) - LEMI*(0.050/0.149);
LY_ECM_1 = lag(LY_ECM, 1);
}

system
{
    "Y" = "DLY";
    "Z" = "Constant",
        "DLY_1", "DLY_2", "DLY_3", "DLY_4",
        LY_ECM_1,
        "DLEMI", "DLEMI_1", "DLEMI_2", "DLEMI_3", "DLEMI_4",
        "DLG", "DLG_1", "DLG_2", "DLG_3", "DLG_4",
        "RLINF", "RLINF_1", "RLINF_2", "RLINF_3", "RLINF_4",
        "LARBDAG", "LARBDAG_1", "DLARBDAG", "DLARBDAG_1",
        "LY-LEMI", "LY-LEMI_1",
        "CSeasonal", "CSeasonal_1", "CSeasonal_2";
}
autometrics(0.01, "IIS", 1);
estimate("OLS", 1996, 3, 2019, 4);

// ECM-equation given by Autometrics
module("PcGive");
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v8.csv");
system
{
    Y = DLY;
    Z = Constant, LY_ECM_1, DLEMI, DLEMI_1, DLEMI_2, DLG_1, DLG_2,
        DLG_3, "I:2008(4)", "I:2011(3)", "I:2011(4)";
}
estimate("OLS", 1996, 3, 2019, 4);


// Static regression with LY, LG and LEMI:
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v8.csv");

system
{
    Y = LY;
    Z = Constant, LG, LEMI;
}
estimate("OLS", 1996, 3, 2022, 3);   // Covid Q's included, as even those quarters shouldn't destroy a long-run relationship if it is there.

// Based on the results, estimating the static ECM-variable
algebra{
LY_static=LY -LG*0.083-LEMI*0.338;
}

```

```
// Estimating marginal models
// Marginal model - DLG
module("PcGive");
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v8.csv");
system
{
    "Y" = "DLG", "DLG_1", "DLG_2", "DLG_3", "DLG_4", "DLG_5", "DLG_6";
    "Z" = "LARBDAG_1", "DLARBDAG", "LARBDAG", "DLARBDAG_1", "CSeasonal", "CSeasonal_1", "CSeasonal_2";
    "U" = "Constant";
}
autometrics(0.01, "IIS", 1);
estimate("OLS", 1996, 3, 2019, 4);

module("PcGive");
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v8.csv");
system
{
    Y = DLG;
    Z = DLG_1, DLG_2, DLG_3, DLG_6, DLARBDAG_1, "I:2000(4)", "I:2009(2)",
        "I:2010(2)", "I:2013(1)";
    U = Constant;
}
estimate("OLS", 1996, 3, 2019, 4);

// Marginal model - DLEMI
module("PcGive");
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v8.csv");
system
{
    "Y" = "DLEMI", "DLEMI_1", "DLEMI_2", "DLEMI_3", "DLEMI_4", "DLEMI_5", "DLEMI_6";
    "Z" = "LARBDAG", "LARBDAG_1", "DLARBDAG", "DLARBDAG_1", "CSeasonal", "CSeasonal_1", "CSeasonal_2";
    "U" = "Constant";
}
autometrics(0.01, "IIS", 1);
estimate("OLS", 1996, 3, 2019, 4); // Changed automatically to 1997(1) - 2019(4) as DLEMI_6 pushes the start of period forward

// Since DLEMI is the time series dating back the shortest, it is adjusted forward to match the rest of
// the estimations, as the four last lags weren't included in the model
module("PcGive");
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v8.csv");
system
{
    Y = DLEMI;
    Z = DLEMI_1, DLEMI_2, "I:1997(1)", "I:1997(2)", "I:1997(3)",
        "I:1998(1)", "I:1998(2)", "I:1998(3)", "I:2001(1)", "I:2003(4)",
        "I:2006(3)", "I:2008(4)";
    U = Constant;
}
estimate("OLS", 1996, 3, 2019, 4);

// Final model system: Conditional + marginal models

```

```
module("PcGive");
package("PcGive", "Multiple-equation");
usedata("GDP_Sweden_v8.csv");
system
{
    Y = DLY, DLG, DLEMI;
    Z = Constant, LY_1, LY_ECM_1,
        DLEMI_1, DLEMI_2, DLG_1, DLG_2, DLG_3,                        // From DLY
        "I:2008(4)", "I:2011(3)", "I:2011(4)",
        DLG_6, DLARBDAG_1, "I:2000(4)", "I:2009(2)",                  // From DLG
        "I:2010(2)", "I:2013(1)",
        "I:1997(1)", "I:1997(2)", "I:1997(3)", "I:1998(1)",           // From DLEMI
        "I:1998(2)", "I:1998(3)", "I:2001(1)", "I:2003(4)", "I:2006(3)";
    I = LY_ECM, LY;
}
model
{
    "DLY" = "Constant", "LY_ECM_1", "DLEMI", "DLEMI_1", "DLEMI_2", "DLG_1", "DLG_2", "DLG_3",
        "I:2008(4)", "I:2011(3)", "I:2011(4)";
    "DLG" = "Constant", DLG_1, DLG_2, DLG_3, DLG_6, DLARBDAG_1, "I:2000(4)",
        "I:2009(2)", "I:2010(2)", "I:2013(1)";
    "DLEMI" = "Constant", DLEMI_1, DLEMI_2, "I:1997(1)", "I:1997(2)", "I:1997(3)",
        "I:1998(1)", "I:1998(2)", "I:1998(3)", "I:2001(1)", "I:2003(4)",
        "I:2006(3)", "I:2008(4)";
    "LY_ECM" = "DLY", "DLG", "DLEMI", "LY_ECM_1";
    "LY" = "DLY", "LY_1";
}
estimate("FIML", 1996, 3, 2019, 4);

// Forecasting
forecast(11); // 2020(1) - 2022(3) // from Covid shock till end of sample

// Estimating the final form equation for DLY
module("PcGive");
package("PcGive", "Single-equation");
usedata("GDP_Sweden_v8.csv");

system
{
    Y = DLY;
    Z = Constant,
        DLY_1,DLY_2,DLY_3, DLY_4,
        DLY_5, DLY_6, DLY_7, DLY_8,
        DLY_9, DLY_10, DLY_11, DLY_12,
        CSeasonal,CSeasonal_1,CSeasonal_2;
}

// Applying Autometrics
autometrics(0.01, "IIS", 1);
estimate("OLS", 1996, 3, 2019, 4);

// Adding the identity equation of LY to the final form equation
module("PcGive");
package("PcGive", "Multiple-equation");
usedata("GDP_Sweden_v8.csv");

system
{
    Y = DLY;
    Z = Constant, LY_1, DLY_2, "I:2008(1)", "I:2008(3)", "I:2008(4)",
```

```
        "I:2009(1)", "I:2010(2)", "I:2011(4)", "I:2012(3)";
    I = LY;
}
model
{   "DLY" = "Constant", "DLY_2", "I:2008(1)", "I:2008(3)", "I:2008(4)",
        "I:2009(1)", "I:2010(2)", "I:2011(4)", "I:2012(3)";
    "LY" = "DLY", "LY_1";
}
estimate("FIML", 1996, 3, 2019, 4);

// Final form equation forecast
forecast(11);
```