Assignment II

Ronan Kenny
Tommy Link
Guus Bouwens
Xavier Raap

March 21, 2023

Data

The dataset contains five quarterly Euro area macro time series: nominal money stock $M3$ (M), GDP deflator, which is a measure for the price level (P), real gross domestic product (Y), nominal housing wealth (WHH), and the annual rate on M3, an interest rate that is called the "own rate" of $M3$ (OR). The order of integration is $I(2)$ for M, P and WHH, and it is $I(1)$ for Y and OR.

In the following sections, we will analyze these time series by setting up models to compute unknown coefficients, discover potential cointegration relations, and investigate the impact of economical change.

Part A

First, we transform the data of M, P, Y, and WHH by taking the natural logarithm. Next, we generate two new variables: 'Real money stock' and 'Growth rate of housing wealth'. We will denote them by $m - p$ and Δwhh, respectively.

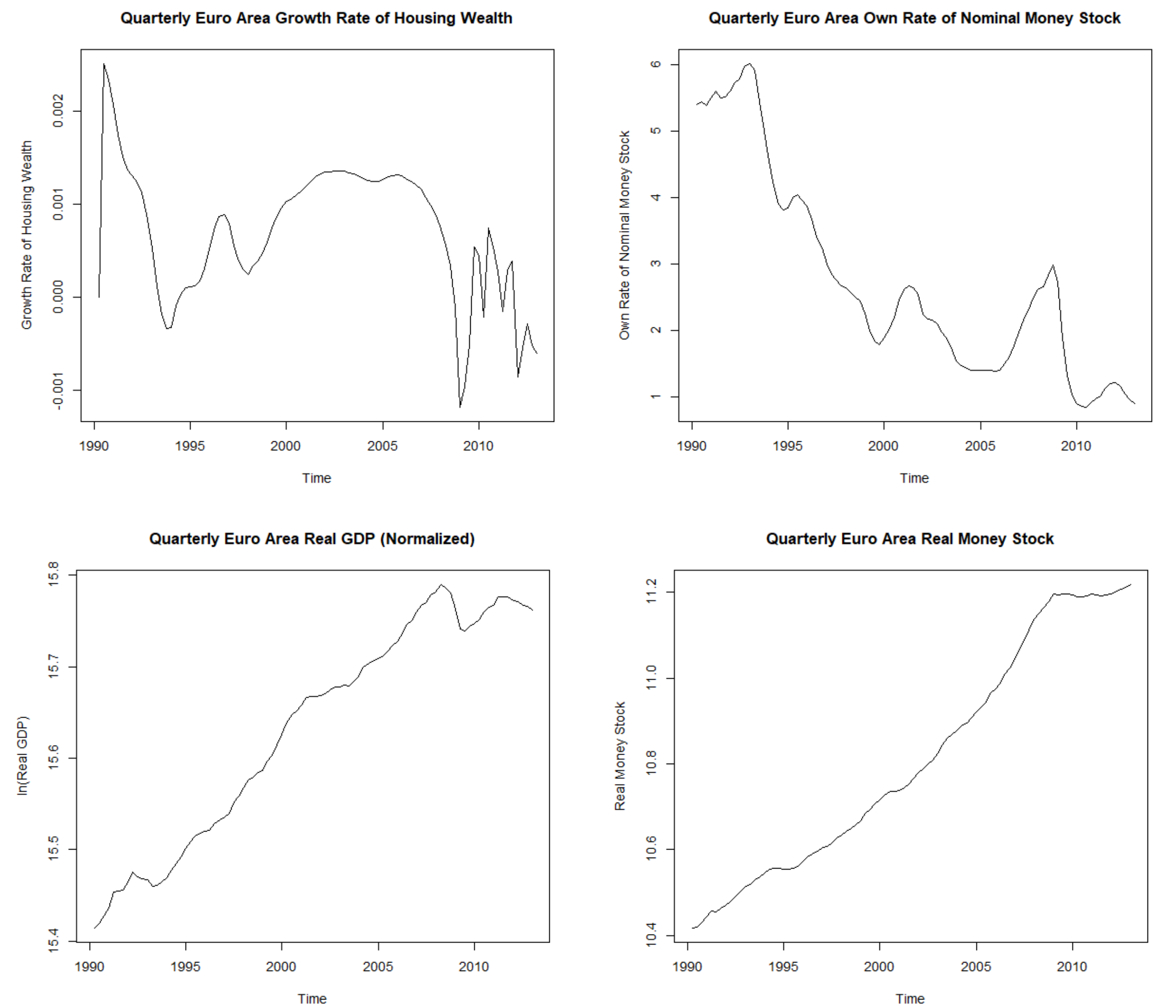

Figure 1: Time series plots of the four $I(1)$ processes

The most apparent feature of the four plots is the crash in the growth rate of housing wealth. This extreme decline from the year 2007 to 2008 was most likely the result of the financial crisis that struck the world around the same time.

Part B

We now turn to the estimation of a vector error correction model ($VECM$) for the three-dimensional system $\mathbf{y} = \begin{bmatrix} m-p \\ y \\ OR \end{bmatrix}$. Before testing for cointegration relations and estimating the unknown coefficients of our VECM, we check which lag order is best to use. In order to do this, we compute the Akaike Information Criterion (AIC) for a $VAR(p)$ model by testing up to two lags. Recall that the formula for the AIC is given by $AIC(p) = \ln(\det(\widehat{\boldsymbol{\Sigma}_{\mathbf{u}}}^{\mathbf{ML}})) + \frac{2}{T}(pK^2 + K)$, where $K = 3$ is the dimension of our $VAR(p)$ model and $\widehat{\boldsymbol{\Sigma}_{\mathbf{u}}}^{\mathbf{ML}}$ is the Maximum-Likelihood estimator of the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{u}}$.

We find values of -25.0 and -26.3 for the AIC(1) and AIC(2), respectively. We, therefore, choose a lag order of $p = 2$ for the VAR model. Since a VECM has a lag order that is one less than its corresponding VAR representation, we conclude that the best lag order to use for our VECM is $p = 1$. The VECM without deterministic terms we set up is:

$$\boldsymbol{\Delta y_t} = \alpha\beta'\mathbf{y_{t-1}} + \Gamma_1\boldsymbol{\Delta y_{t-1}} + \mathbf{u_t} \tag{1}$$

The $(K \times r)$ matrices α and β are the loading and cointegration matrix, respectively. The $(r \times 1)$ vector $\beta'\mathbf{y_{t-1}}$ contains all the r cointegration relations. We maximize the log-likelihood function in order to find the estimates for α, β, Γ_1 and Σ_u (see slides 26-27 of week 5 for explicit formulae). We obtain the following estimates:

$$\hat{\alpha} = \begin{bmatrix} 0.01 \\ 0.01 \\ -0.15 \end{bmatrix} \quad \hat{\beta} = \begin{bmatrix} -1.00 \\ 0.72 \\ -0.06 \end{bmatrix} \quad \hat{\Gamma}_1 = \begin{bmatrix} 0.67 & -0.18 & 0.01 \\ -0.07 & 0.58 & -0.01 \\ -2.06 & 12.00 & 0.62 \end{bmatrix} \quad \widehat{\Sigma_u} = \begin{bmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.01 \end{bmatrix}$$

Next, we perform a sequence of hypothesis tests to find the number of cointegration relations: $H_0 = rk(\Pi) = r_0$ vs $H_1 = rk(\Pi) > r_0$, where $r_0 = 0$ and increments by one each time we perform a new test until $r_0 = K - 1$. We use the vector of eigenvalues $\lambda = \begin{bmatrix} 0.29775787 & 0.04893553 & 0.03946885 \end{bmatrix}'$ to compute the test statistic (see slide 31 of week 5). The trace test results for cointegration are as follows:

r	trace	p-value$_{trace}$	eigen	p-value$_{eigen}$
0	39.953	< 0.001	31.813	< 0.001
1	8.140	0.228	4.516	0.553
2	3.624	0.066	3.624	0.068

Table 1: Johansen's test results

The output provides two tests:

- Trace - tests the null hypothesis of $rk(\Pi) = r$ against the alternative: $r < rk(\Pi) \leq K$
- Eigenvalue - tests the null hypothesis of $rk(\Pi) = r$ against the alternative: $rk(\Pi) = r + 1$

We are interested in the latter of these tests. Given a significance level of $\alpha = 1\%$ for the cointegration test decision, we can see that there is one cointegrating relationship ($r = 1$) present in the system, as the null hypothesis of $r = 0$ cointegrating relationships is rejected at a 1% significance level, while the null hypotheses of $r = 1$ and $r = 2$ are not rejected.

We will now interpret the obtained results. We know that the decomposition $\Pi = \alpha\beta'$ is not unique. Hence, we must impose the restriction that the first $r \times r$ block of our $K \times r$ matrix β is equal to $\mathbb{I}_r$. In our case $r = 1$ and we can simply multiply both $\hat{\alpha}$ and $\hat{\beta}$ by -1 to obtain the desired form. Then $\hat{\beta} = \begin{bmatrix} 1.000 & -0.723 & 0.057 \end{bmatrix}'$ and the estimated loading matrix $\hat{\alpha}$ has the estimated coefficients:

Variable	Coefficient
m.minus.p	-0.012 (0.003)
y	-0.006 (0.003)
OR	0.148 (0.066)

Table 2: Loading matrix α

Using a 10% significance level, we can see that all the coefficients for the loading matrix are statistically significant. The α-coefficients represent the speed at which the variables in the system adjust to deviations from the long-run equilibrium. A negative coefficient indicates that deviations from the equilibrium will lead to a decrease in the variable in the subsequent period, while a positive coefficient suggests that deviations will lead to an increase in the variable. This is of course only the case if the corresponding coefficient in the cointegration matrix β is positive. Otherwise, we will have a doubly negative effect and hence a positive effect. It is therefore important to look at the found cointegration relation as well: $\hat{\beta}'\mathbf{y_{t-1}} = RMS_{t-1} - 0.723\Delta RGDP_{t-1} + 0.057OR_{t-1}$. So a positive economic shock in either the real money stock (RMS), growth of real GDP (ΔGDP) or annual rate on M3 will increase $\hat{\beta}'\mathbf{y_{t-1}}$ in the first and third case and decrease in the second case due to the coefficient of $\hat{\beta}$. It is therefore important to keep these effects in mind, as we will now consider the α-coefficients and their effect on $\hat{\beta}'\mathbf{y_{t-1}}$. For the real money stock ($m-p$ variable), the α-coefficient is significantly negative (-0.012) at a 10% significance level. This implies that deviations from the long-run equilibrium have a notable but small negative impact on this variable in the short-run. For the growth of the real GDP (y variable), the α-coefficient is also significantly negative (-0.006). This suggests that deviations from the long-run equilibrium also have limited diminishing effect on the growth of real GDP in the short-run. For the OR variable (the annual rate on M3), however, the α-coefficient is significantly positive (-0.148). This means that if the annual rate on M3 deviates from its long-run equilibrium, it will adjust in the same direction in the short-run. In other words, if the annual rate on M3 is higher than its long-run equilibrium, it will increase in the following period, and vice versa. In summary, our VECM results indicate that there is one cointegrating relationship in the system, namely that $\hat{\beta}'\mathbf{y_{t-1}} = RMS_{t-1} - 0.723\Delta RGDP_{t-1} + 0.057OR_{t-1}$. Furthermore, all variables are affected by deviations from the long-run equilibrium in the short-run.

Part C

To extend the model by the growth rate of housing wealth (Δwhh) and exclude all observations from 2007Q3 onward, we first need to adjust the dataset. This creates the system $\mathbf{y} = \begin{bmatrix} m-p \\ y \\ OR \\ \Delta whh \end{bmatrix}$. We will then modify the VECM estimation function to include an overall (unrestricted) intercept and a linear trend that is restricted to the cointegration relations. This leads to the following VECM: with

$$\boldsymbol{\Delta y_t} = \alpha\beta^{o\prime}\mathbf{y^o_{t-1}} + \Gamma_1\boldsymbol{\Delta y_{t-1}} + \mathbf{u_t} \tag{2}$$

In the equation above, $y^o_{t-1} = \begin{pmatrix} y_{t-1} \\ 1 \end{pmatrix}$. The term $\beta^{o\prime} = (\beta' : -\beta'\mu_0)$ includes the cointegration matrix and the constant. This incorporates the constant term in the matrix Δx_t, a trend in matrix y_{t-1}. The variables and the procedure of estimation of the VECM are the same as described in the section under Equation 1 in part B. We obtain the following results:

$$\hat{\alpha} = \begin{bmatrix} 0.5088 \\ 0.0033 \\ 0.0015 \\ -0.0297 \\ -0.00003 \end{bmatrix} \quad \hat{\beta^{o\prime}} = \begin{bmatrix} 0.0483 \\ 0.0958 \\ -0.0021 \\ -1.00 \\ -0.001 \end{bmatrix} \quad \hat{\Gamma_1} = \begin{bmatrix} -0.0137 & -0.0713 & -0.0008 & 0.0722 \\ 0.5604 & -0.0336 & 0.0075 & 1.5527 \\ -0.0151 & 0.4284 & -0.003 & 1.3582 \\ 1.3236 & 7.1148 & 0.7577 & -5.9816 \\ -0.001 & 0.0075 & -0.0002 & 0.0713 \end{bmatrix}$$

$$\widehat{\Sigma_u} = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.01 & 0.0001 & 0.0000 \\ 0.0000 & 0.0000 & 0.0001 & 0.0085 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}$$

The Δwhh (the last row of each of the estimates) coefficients represent the short-run dynamics of the VECM model. The coefficients in the alpha matrix show the speed of adjustment to the long-run equilibrium, while the coefficients in the Γ_1 matrix show the short-run impact of each variable on other variables in the system. We examine the eigenvalues and the Johansen test results to see if there's evidence for cointegration between the variables in the system. The eigenvalues are $\lambda = \begin{bmatrix} 0.563 & 0.459 & 0.207 & 0.091 & 0.00 \end{bmatrix}'$. The Johansen test results show the maximal eigenvalue statistic (λ_{max}) with a linear trend in cointegration. The eigenvalues are sorted in descending order, and their magnitudes indicate the strength of the cointegration relationships. The null hypothesis is that there are at most r cointegrating relationships. Below is the table of test statistics and critical values of the test:

		Critical value		
Null hypothesis	Test statistic	10%	5%	1%
r = 0	57.10	29.12	31.46	36.65
r $\leq$ 1	42.34	23.11	25.54	30.34
r $\leq$ 2	16.02	16.85	18.96	23.65

Table 3: Johansen's test results for VECM with linear trend and intercept

We start with the null hypothesis r = 0 and compare the test statistic (57.10) to the critical values at different significance levels (10%, 5%, and 1%). Since the test statistic is greater than the critical values, we reject the null hypothesis. Next, we test the null hypothesis r $\leq$ 1. The test statistic (42.34) is again greater than the critical values, so we reject the null hypothesis. When we test the null hypothesis r $\leq$ 2, the test statistic (16.02) is smaller than the critical values for all three significance levels, which indicates that we fail to reject the null hypothesis at a 5% (even at 1%) significance level. Based on the Johansen test results, we thus find evidence for two cointegrating relationships at a 5% significance level. This suggests that there are two long-run equilibrium relationships among the variables in the model. Since we only found evidence for one cointegration relation in the previous VECM, extending the model by the growth rate of housing wealth could imply that this economic factor also plays an essential role in the system.

Part D

To investigate the impact of the financial crisis on the stability of the model coefficients, we should re-estimate the VECM model using all available observations in the dataset, including those after the financial crisis. After re-estimating the VECM model using all available observations, i.e. for the system $\mathbf{y} = \begin{bmatrix} m-p \\ y \\ OR \\ \Delta whh \end{bmatrix}$, we examine the summary output to evaluate the stability of the model coefficients. We compare the estimated coefficients and their statistical significance to those obtained in our previous model. If the coefficients and their significance levels have changed considerably, this could suggest that the financial crisis has had an impact on the stability of the model coefficients. If the coefficients remain relatively stable and maintain their significance levels, it could indicate that the model is robust and stable even when including the financial crisis period in the analysis. Here are the coefficients:

$$\hat{\alpha} = \begin{bmatrix} 0.8613 \\ -0.0001 \\ 0.0033 \\ -0.0318 \\ -0.0009 \end{bmatrix} \quad \hat{\beta} = \begin{bmatrix} -1.00 \\ 0.01 \\ 0.0001 \\ 0.0685 \\ -0.0004 \end{bmatrix} \quad \hat{\Gamma_1} = \begin{bmatrix} 0.4212 & -0.1001 & -0.0013 & 0.0178 \\ -0.0508 & -0.0038 & -0.0007 & 0.0089 \\ 1.7958 & 0.6530 & 0.0073 & -0.0941 \\ 42.5088 & -1.8647 & 0.6678 & 9.2828 \\ 1.7285 & -0.0153 & -0.0068 & 0.6590 \end{bmatrix}$$

$$\widehat{\Sigma_u} = \begin{bmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.0001 & 0.0111 & 0.0002 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.0002 & 0.00 \end{bmatrix}$$

The eigenvalues are $\lambda = \begin{bmatrix} 1.00 & 0.266 & 0.165 & 0.118 & 0.097 \end{bmatrix}'$

		Critical value		
Null hypothesis	Test statistic	10%	5%	1%
r = 0	31.74	29.12	31.46	36.65
r ≤ 1	16.45	23.11	25.54	30.34

Table 4: Johansen's test results for VECM with linear trend and intercept (all data)

Similar as to what was done in part C, we perform the Johansen test on the system based off the table of test statistics above. For null hypothesis r = 0, the test statistic (31.74) is greater than the critical values at a 10% and 5% significance level, but is less than the critical value at a 1% significance level. Therefore, we reject the null hypothesis if we have a 5% or 10% significance level, and fail to reject it if we have a 1% significance level. For null hypothesis r ≤ 1, we have that the test statistic (16.45) is less than all the critical values at the 10%, 5%, and 1% significance levels. We will fail to reject the null hypothesis for all significance levels. The Johansen test result therefore tells us that there is one cointegrating relationship if we use a 5% significance level.

We can observe from the code that the eigenvectors and their corresponding cointegration relationships appear to be relatively stable, with minor changes compared to the previous model. The adjustment coefficients in the alpha matrix show some variations, but overall, they are still relatively stable. The phi matrix, which represents the short-run dynamics, has also experienced some changes, but the coefficients remain in a reasonable range. The omega matrix, representing the variance-covariance matrix of the residuals, has also undergone some changes but still maintains a similar structure to the previous model.

In conclusion, based on the re-estimated VECM model using all available observations, the stability of the model coefficients has been affected to some extent by the financial crisis. However, the model's overall structure remains relatively stable, and the Johansen procedure indicates the presence of one cointegrating relationship among the variables. This suggests that the VECM model is still robust enough to analyze the long-run relationships between the variables, even when including the financial crisis period in the analysis.

Appendix

Reading in data

```
rm(list=ls())
library(readr)
data <- read.csv("data_assignment2.csv", dec=",", sep=";")
```

Code for part A

```
data.log <- log(data[,2:5])
colnames(data.log) <- c("m","y","whh","p")
data.log$m.minus.p <- (data.log$m - data.log$p)

data.log$delta.whh[1] <- 0
for(i in 2:(nrow(data.log))){
  data.log$delta.whh[i] <-  (data.log$whh[i] - data.log$whh[i-1]) / data.log$whh[i-1]
}
m.minus.p.data <- data.log[5]
y.data <- data.log[2]
delta.whh.data <- data.log[6]
or.data <- data[6]
plot(ts(m.minus.p.data, frequency = 4, start=c(1990,2)), main="Quarterly Euro Area Real Money Stock",
plot(ts(y.data, frequency = 4, start=c(1990,2)), main="Quarterly Euro Area Real GDP (Normalized)", yl
plot(ts(delta.whh.data, frequency = 4, start=c(1990,2)), main="Quarterly Euro Area Growth Rate of Hou
```

```
plot(ts(or.data, frequency = 4, start=c(1990,2)), main="Quarterly Euro Area Own Rate of Nominal Money
```

Code for part B

```
library(vars)
library(geigen)
data$delta.whh <- data.log$delta.whh
data$m.minus.p <- (data.log$m - data.log$p)
data$y <- data.log$y
vecm.data <- data[c(7,8,6)]
ts.vecm.data <- ts(vecm.data, frequency=4,start=c(1990,1))
VARselect(ts.vecm.data, lag.max = 2)

p <- 1
r <- 1

build_dX <- function(dataset, p){
  return(dataset[(p+2):nrow(dataset),] - dataset[(p+1):(nrow(dataset)-1),])
}

build_Xlag <- function(dataset, p){
  return(dataset[(p+1):(nrow(dataset)-p),])
}

build_W <- function(dataset, p){
  dX_full <- t(dataset[2:(nrow(dataset)-1),] - dataset[1:(nrow(dataset)-2),])
  W <- matrix(c(dX_full[, p:1]), 1)
  for (t in (p+1):ncol(dX_full)){
  W <- rbind(W, matrix(c(dX_full[,(t-p+1):t]), 1))
  }
  return(W)
}

build_Mw <- function(W){
  Mw <- (W %*% solve(t(W) %*% W) %*% t(W))
  return(diag(1, nrow(Mw)) - Mw)
}

build_S00 <- function(Mw, dX){
  return((t(dX) %*% Mw %*% dX) / nrow(dX))
}

build_S11 <- function(Mw, Xlag){
  return((t(Xlag) %*% Mw %*% Xlag) / nrow(Xlag))
}

build_S01 <- function(Mw, Xlag, dX){
  return((t(dX) %*% Mw %*% Xlag) / nrow(Xlag))
}

est_alpha <- function(beta, S01, S11){
  return(S01 %*% beta %*% solve(t(beta) %*% S11 %*% beta))
}

est_phi <- function(beta, dX, Xlag, W, alpha){
  return(t(dX - (Xlag %*% beta %*% t(alpha))) %*% W %*% solve(t(W) %*% W))
}
```

```
est_omega <- function(beta, S00, S01, S11){
  S00 - (S01 %*% beta %*% solve(t(beta) %*% S11 %*% beta) %*% t(beta) %*% t(S01))
}

est_beta <- function(S11, S00, S01, r){
  A <- (t(S01) %*% solve(S00)) %*% (S01)
  B <- S11
  eigenvalues <- geigen(A, B, symmetric=FALSE)
  beta <- eigenvalues$vectors[,order(-eigenvalues$values)]
  beta <- matrix(beta[, 1:r], ncol=r)
  lambdas <- eigenvalues$values[order(-eigenvalues$values)]
return(list(beta, lambdas))
}

est_VECM <- function(dataset, p=1, r=1){
  dX <- build_dX(dataset, p)
  Xlag <- build_Xlag(dataset, p)
  W <- build_W(dataset, p)
  Mw <- build_Mw(W)
  S00 <- build_S00(Mw, dX)
  S11 <- build_S11(Mw, Xlag)
  S01 <- build_S01(Mw, Xlag, dX)
  eigenvalues_results <- est_beta(S11, S00, S01, r=r)
  beta <- eigenvalues_results[[1]]
  lambdas <- eigenvalues_results[[2]]
  alpha <- est_alpha(beta, S01, S11)
  phi <- est_phi(beta, dX, Xlag, W, alpha)
  omega <- est_omega(beta, S00, S01, S11)
return(list(beta=beta, alpha=alpha, phi=phi, omega=omega,
            lambdas=lambdas, S00=S00, S11=S11))
}

VECM_results <- est_VECM(ts.vecm.data, p, r)
VECM_results
LRstat <- function(lambdas, n=nrow(ts.vecm.data - 2)){
  -n*sum(log(1 - lambdas))
}
lambdas <- est_VECM(ts.vecm.data, p, r)$lambdas
LRstat(lambdas)

summary(rank.test(VECM(ts.vecm.data,lag=1,r=1,include='none',estim='ML',LRinclude='none'),cval=0.01))
```

Code for part C

```
# Find the index corresponding to 2007Q2
index_2007Q2 <- which(data$X == "2007-4")

# Create the restricted dataset
data_restricted <- data[1:(index_2007Q2 - 1), ]
data_restricted$log_delta_whh <- data.log$delta.whh[1:nrow(data_restricted)]
vecm_data_restricted <- data_restricted[c("m.minus.p", "y", "OR", "log_delta_whh")]

# Convert the dataset to a time series object
ts_vecm_data_restricted <- ts(vecm_data_restricted, frequency = 4, start = c(1990, 1))

est_VECM_trend <- function(dataset, p = 1, r = 1) {
```

```
  dX <- build_dX(dataset, p)
  Xlag <- build_Xlag(dataset, p)
  W <- build_W(dataset, p)

  # Add constant term to dX
  dX_with_const <- cbind(1, dX)

  # Add linear trend to Xlag
  trend <- 1:nrow(Xlag)
  Xlag_with_trend <- cbind(Xlag, trend)

  Mw <- build_Mw(W)
  S00 <- build_S00(Mw, dX_with_const)
  S11 <- build_S11(Mw, Xlag_with_trend)
  S01 <- build_S01(Mw, Xlag_with_trend, dX_with_const)

  eigenvalues_results <- est_beta(S11, S00, S01, r = r)
  beta <- eigenvalues_results[[1]]
  lambdas <- eigenvalues_results[[2]]
  alpha <- est_alpha(beta, S01, S11)
  phi <- est_phi(beta, dX_with_const, Xlag_with_trend, W, alpha)
  omega <- est_omega(beta, S00, S01, S11)

  return(
    list(
      beta = beta,
      alpha = alpha,
      phi = phi,
      omega = omega,
      lambdas = lambdas,
      S00 = S00,
      S11 = S11
    )
  )
}

# Estimate VECM with the modified function
VECM_trend_results <- est_VECM_trend(ts_vecm_data_restricted, p, r)

# Run the VECM using the vars package
restricted_vecm <- ca.jo(
    ts_vecm_data_restricted,
    type = "eigen",
    ecdet = "trend",
    K = 2, # Set the lag order to 2
    spec = "transitory"
)

summary(restricted_vecm)
```

Code for part D

```
# Convert the full dataset to a time series object
vecm.data.partd <- ts(data[c(7,8,6,9)], frequency = 4, start = c(1990,1))

# Estimate VECM with the modified function
VECM_trend_results_full <- est_VECM_trend(vecm.data.partd, p, r)
```

```
VECM_trend_results_full

# Run the VECM using the vars package
full_vecm_d <- ca.jo(
    vecm.data.partd,
    type = "eigen",
    ecdet = "trend",
    K = 2, # Set the lag order to 2
    spec = "transitory"
)

summary(full_vecm_d)
```