GROUP #1
Roberto Vacante 0001131429 – *Francesco Colzi* 0001127782
Daniel Vito Lobasso 0001125016 – *Danny Prest* 0001140688 – *Ciro Spallitta* 0001140660

PC EXAM

(1) Introduction of the work. First indicate the labels of your variables used in Gretl/Stata. Then, make a brief description/commentary of the long-term relationship studied, indicating and justifying from the economic point of view the expected direction of causality, and the sign of the parameters.

The specification of the model is given by the following equation:

$$\Delta log(dvcft)_t = \beta_0 + \beta_1 \Delta log(ulc)_t + \beta_2 \Delta log(dmbs)_t + \beta_3 log(qdisi)_t + \varepsilon_t$$

In this framework, quarterly data spanning from 1960 to 2022 are used for the variables of interest, namely: the change in the deflator of value added (*d_l_dvcft*), the change in unit labour cost (*d_l_ulc*), the change in the deflator of imports (*d_l_dmbs*), and the unemployment rate (*l_qdisi*). First-differences of $log(dvcft)$, $log(ulc)$ and $log(dmbs)$ are treated as level variables of inflation and growth rates. The table below offers further clarifications on the labels used for the aforementioned variables.

Variables	vars_name	Labels
$\Delta log(dvcft)$	*d_l_dvcft*	Quarterly (internal) inflation rate of value added.
$\Delta log(ulc)$	*d_l_ulc*	Quarterly inflation rate of wages per unit of output (wage inflation).
$\Delta log(dmbs)$	*d_l_dmbs*	Quarterly (external) inflation rate of import.
$log(qdisi)$	*l_qdisi*	Logarithm of unemployment rate.

In this model, a Phillips Curve relationship is implicit in explaining domestic inflation. Indeed, it suggests that changes in internal inflation are influenced by changes in wage inflation, import inflation and the unemployment rate, reflecting the traditional trade-off between inflation and unemployment.

In particular, we expect that:

- An increase in unit labour costs should lead to higher production costs for businesses. Since companies often pass these increased costs onto consumers, it's anticipated that there will be a direct and positive relationship with the deflator of value added. Thus, β_1 is expected to be positive, indicating that as labour becomes more expensive, the general price level of domestically produced goods and services increases.

- An increase in import prices can raise the overall cost of inputs for domestic production, especially for economies that rely heavily on imported goods and services. This scenario also contributes to inflationary pressures, suggesting a positive β_2. It reflects the transmission of external price shocks to the domestic economy, contributing to the cost-push inflation where higher import prices lead to increased prices of domestically produced goods and services.

- The unemployment rate's inclusion, particularly in logarithmic form, reflects the theoretical underpinnings of the Phillips curve, which traditionally posits an inverse relationship between unemployment and inflation. In the short term, lower unemployment is associated with upward pressure on wages due to higher demand for labour. This, in turn, drives up production costs, resulting in higher prices for goods and services. However, the sign of β_3 in the long-term context could vary. Contemporary economic studies frequently underscore the distinction drawn by recent Phillips curve formulations between a short-term and a long-term trajectory, the latter often conceptualised as the Non-Accelerating Inflation Rate of Unemployment (NAIRU). This distinction particularly emphasises the potential erosion of the inflation-unemployment trade-off over the long term. Such weakening of the relationship can be attributed to an array of factors, including globalisation, technological progress, and the evolution of monetary policy frameworks.

(2) Summarise what are the main findings of your work. Even with respect to the ex ante predictions (see the point above).

This study aims to clarify the determinants of domestic inflation, focusing on the roles of imported inflation, wage inflation, and the unemployment rate. Initial univariate analyses distinguished between stationary (imported inflation and wage inflation) and non-stationary variables (domestic inflation and the unemployment rate), setting the stage for a comprehensive multivariate examination. Through the application of Autoregressive Distributed Lag (ARDL) and Error Correction Models (ECM), we investigated the long-term relations among these variables. Our findings reveal a consistent speed of adjustment across models, indicating a gradual process of adjustment for domestic inflation toward its long-term equilibrium. Notably, discrepancies in the long-run parameter estimates were minimal, with imported inflation presenting the sole exception, potentially attributable to its stationarity.

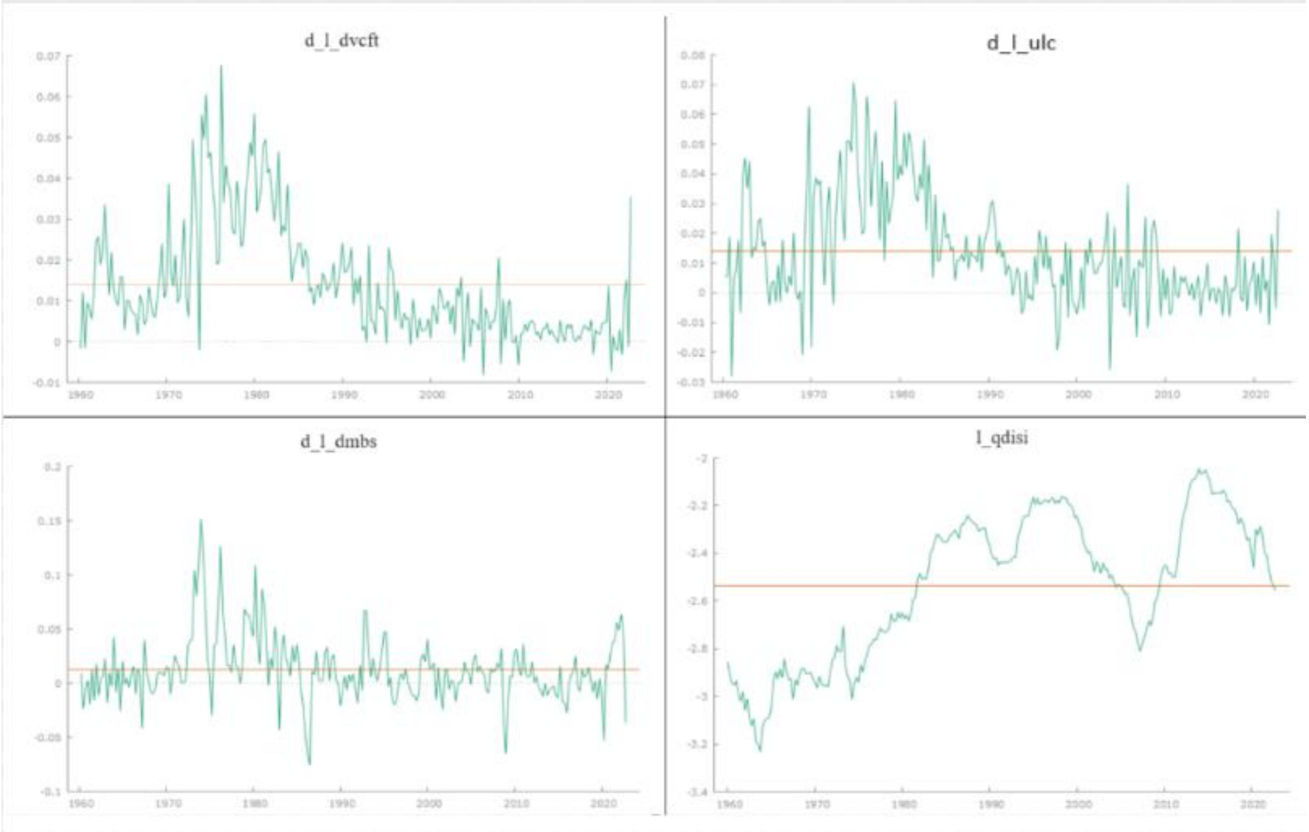
To clarify the nature of the long-run relationship between the variables, we conducted further analysis through the Engle-Granger procedure and Johansen cointegration approach to identify the most realistic scenario for a long run relationship between the variables, considering a rank of the matrix equal to 2. Despite the initial univariate analysis could lead us to believe that the long run relationship could only be explained by the I(0) variables (*d_l_ulc* and *d_l_dmbs*) being cointegrated with themselves, we found evidence that wage inflation may actually be non- stationary. Indeed, we believe that *d_l_ulc* is cointegrated with domestic inflation *d_l_dvcft*. In light of this result, the rank equal to 2 of the matrix can be explained by this cointegration and by the stationarity of imported inflation, which is auto-cointegrated.

The vector error correction model (VECM) makes the multi-equation results of the long run parameter preferred over single-equation models, since the latter do not satisfy the hypotheses concerning the weak exogeneity required by PSS and EG approaches.
Concerning the multi-equation results, they exhibit a positive long-run relationship between wage inflation and domestic inflation, as well as for single models, but with a higher value. Moreover, according to the unemployment rate, our findings suggest a lack of substantial causation with the dependent variable. Differently, imported inflation points out a cointegrating relationship with itself.

In conclusion, our analysis suggests the existence of an endogenous long-term cointegrated relationship between value-added inflation and unit labour cost inflation, which is also economically explainable. This simultaneous causality implies that an increase in unit labour costs prompts firms to raise product prices to maintain profitability. At the same time, conventional macroeconomic models indicate that inflation (and expected inflation) plays a significant role in wage dynamics, contributing to the wage-price spiral phenomenon.
On the other hand, import inflation seems relevant only in the short term, as evidenced by its impact during the recent 2022 energy crisis. Plausible explanations include economic actors reacting to shocks through policy adjustments (monetary, fiscal, or supply-side), or shocks reverting on their own. Therefore, over the long term, the influence of imports may diminish, given the ongoing evolution of the broader economic context.
Finally, our analysis indicates that the unemployment rate does not exhibit a long-term relationship with value-added inflation, as evidenced by nonsignificant estimates and recent studies on the Phillips curve, particularly regarding the absence of a trade-off between unemployment and inflation (NAIRU).

(3) Univariate preliminary analysis, 1/3. Plot all the variables that are listed in the long run of your specification, and their first differences. Conclude with a short comment of the pattern of the components of your relationship and their first differences. Do they look stationary? How do you expect correlograms and ADF tests will be in the light of these plots?

Plot of levels variables



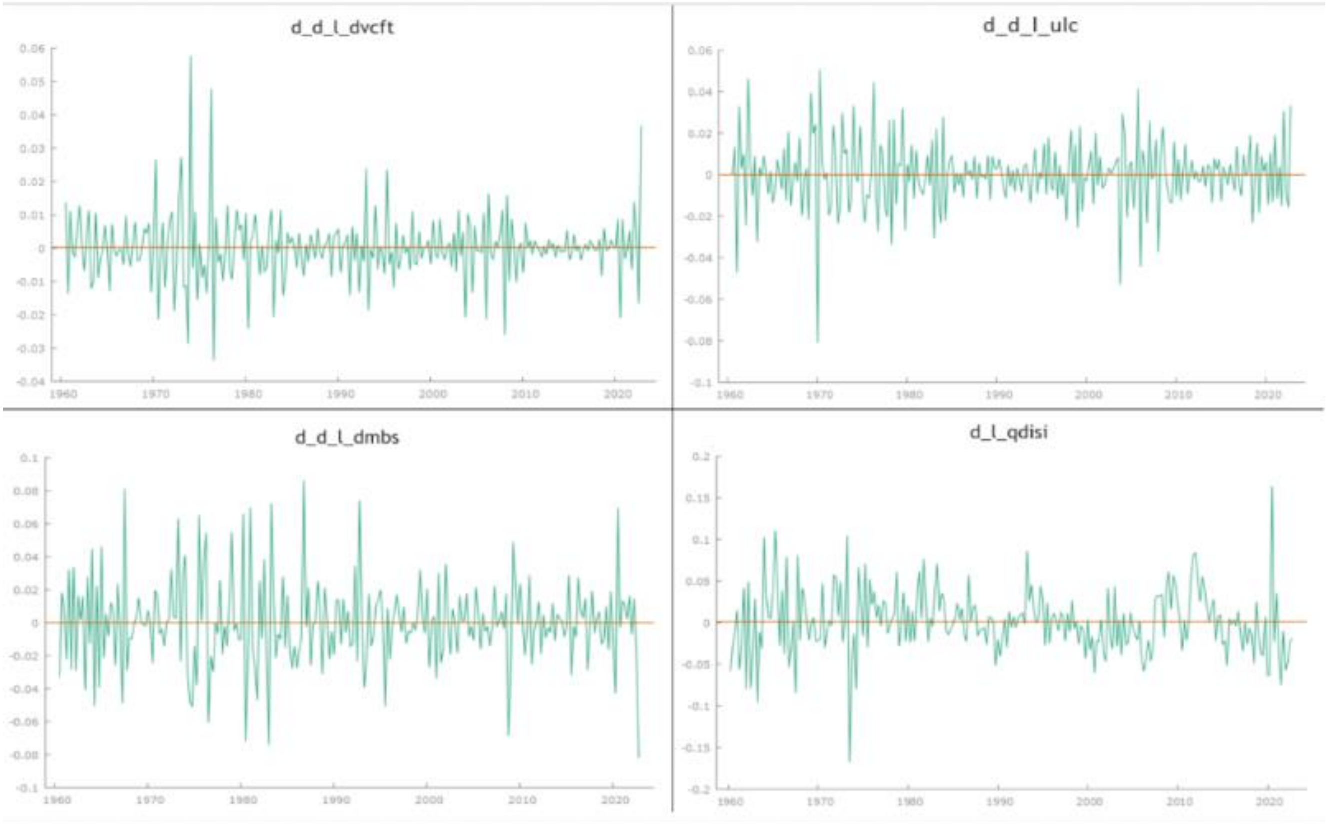
From an economic point of view, inflation plots show a growth in prices between the 1970s and 1980s, probably due to the oil shocks that have hit many countries during that period.
A sharp increase in unemployment can be observed in the same years, probably linked to restrictive monetary policies that led to a recession and therefore worsened the job market situation. However, the post-2008 situation is different, showing a positive trend for unemployment (presumably due to the economic and financial crisis), which is not supported by a strong change in inflation in any of the cases analysed.

Judging from the Level plots shown above, only $\Delta log(dmbs)$ looks stationary. Despite some visibly high shocks, the values oscillate around the mean for most of the series. The other variables clearly present some persistence, which seems to be the highest for $log(qdisi)$.

Furthermore, $\Delta log(dvcft)$ and $\Delta log(ulc)$ exhibit a similar pattern throughout the series, which could possibly reflect the existence of a positive relationship between the two variables, as stated earlier in the introduction. As the level variables are ratios, none of them seem to present a clear trend, even those that seem nonstationary.

In light of these considerations, we can reasonably expect to see high coefficients of autocorrelation for many lags and non-significant ADF tests for most of these variables.

Plot of differentiating variables

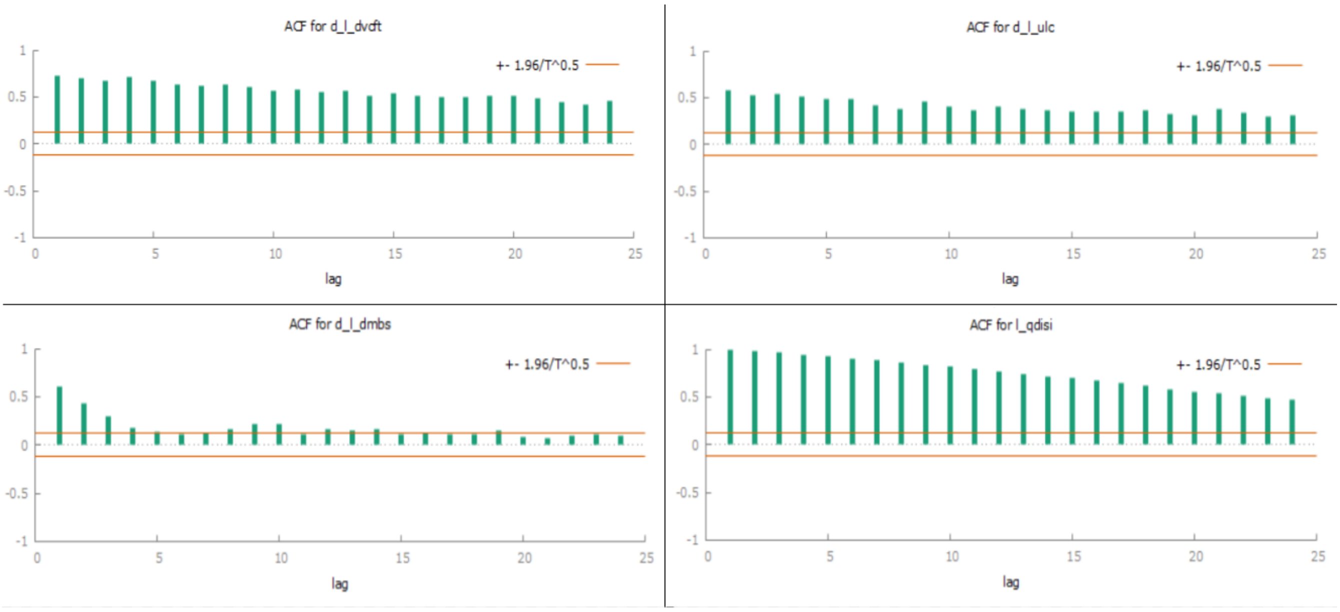


Unsurprisingly, as we compute the first differences of the level variables, the plots of the series become stationary. Moreover, their mean drops to zero, reflecting the absence of a trend. We can reasonably expect that the first differenced variables are integrated of order 0 as they clearly show no signs of persistence.

In light of these considerations, we can reasonably expect to see correlograms with no-significant coefficients of autocorrelation (most of them within the interval confidence) and significant ADF tests for these variables.

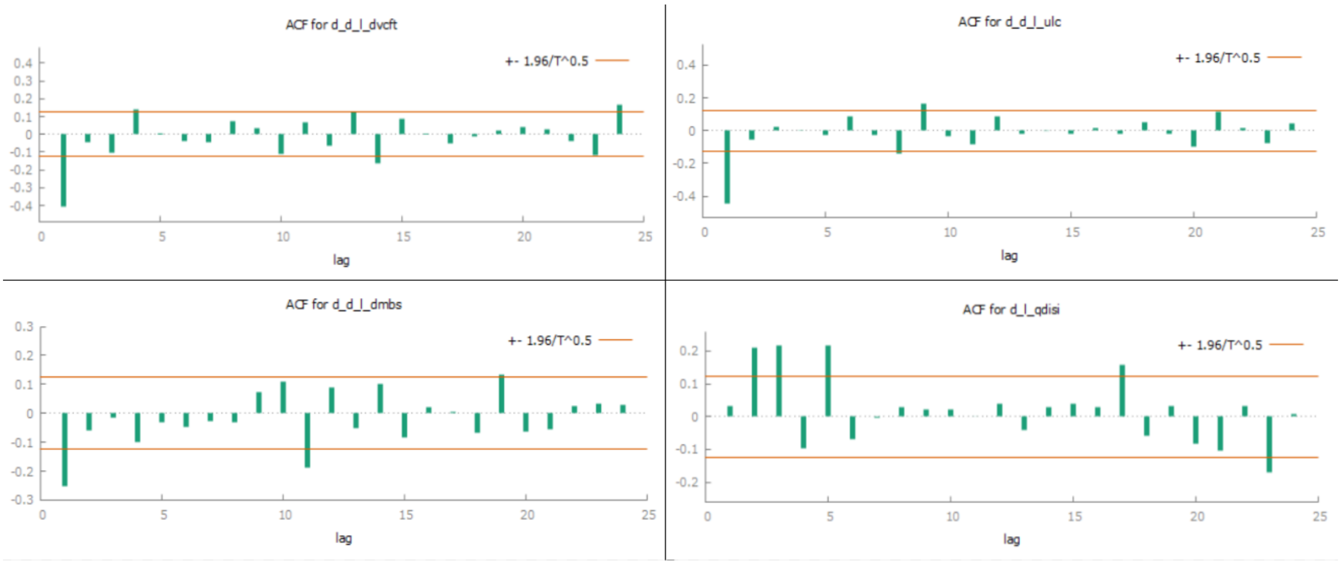
(4) Univariate preliminary analysis, 2/3. Show all the correlograms corresponding to the plots above and assess if these outcomes are in line with the above.

Correlograms of levels variables



As expected, the autocorrelation plots suggest a high level of persistence for most of the variables. In particular, $\Delta log(dvcft)$, $\Delta log(ulc)$ and $log(qdisi)$ exhibit serial correlation. In line with the previous analysis, $log(qdisi)$ exhibits the highest autocorrelation, which is very close to one in the first lag.
The level of autocorrelation of $\Delta log(dmbs)$ (expressing the import inflation), drops inside the confidence interval from the fourth lag onward, confirming that the variable is probably stationary.

Correlograms of differentiating variables



As we take the first differences, the autocorrelation plots show an oscillating behaviour, mostly inside the confidence interval, which seems to be in line with what is seen above. The unusual plot of *d_l_qdisi* suggests serial correlation for some of the first lags, but the absence of regular peaks at the 4th, 8th, 12th lag seem to suggest a lack of seasonality across the series.

(5) Univariate preliminary analysis, 3/3. ADF tests outcomes of both levels and first differences. Results must be commented (be consistent with your claims above) from a summary table, where all relevant statistics of the test are reported.

Variables	Deterministic components*	Augmentation	T-statistic	P-value	Integration
Levels					
d_l_dcvft	c	3	-2.291	0.1750	I(1)
d_l_ulc	c	3	-3.353	0.0127	I(0)
d_l_dmbs	c	0	-7.592	0.0000**	I(0)
l_qdisi	c	4	-1.709	0.4269	I(1)
First difference					
d_d_l_dcvft	c	2	-15.78	0.0000**	I(0)
d_d_l_ulc	c	2	-14.72	0.0000**	I(0)
d_d_l_dmbs	c	0	-20.16	0.0000**	I(0)
d_l_qdisi	c	3	-6.568	0.0000**	I(0)

* c = constant. This choice is justified by the fact that the explanatory variables represent rates (no trends).

** When the value comes to 0.0000 it means that the p-value is tending to zero (as approximation, we consider only the first four values after the dot).

The table shown above summarises the outcomes of the ADF tests on both level and first differenced variables. Generally, series with quarterly data could be treated as AR(4) processes, suggesting an ADF(3) test as the most suitable. However, this was not the case for all the variables. As indicated by different information criteria, no augmentation was needed for *d_l_dmbs*, suggesting the absence of persistence in the series, which is in line with our previous analyses. Moreover, for *l_qdisi*, the results suggest an ADF(4) test, which could be due to the higher persistence we previously observed.

Moving to the outcomes, the coefficients for *d_l_dcvft* and *l_qdisi* are non-significant, meaning that they have a unit root and thus are I(1). Contrary to our initial expectations, the coefficient for *d_l_ulc* is significant at the 5% level, indicating that, despite some visible degrees of persistence in the plots, the variable has no unit root and thus it is I(0).

After first differencing, the variables that were non-stationary at the levels become I(0). For the sake of completeness, we carried out ADF tests even for the variables that were I(0) at the levels. Differencing these variables further can lead to distorted or poorly informative results. Indeed, information criteria suggest using high augmentations to carry out these tests, such as ADF(5) and ADF(6).

For this reason, we wanted to continue following the reasoning made for the integrated variables by applying an augmentation of one lower order than the levels, being variables in differences. For *d_d_l_dmbs*, not being possible to have a negative augmentation, we maintained 0.

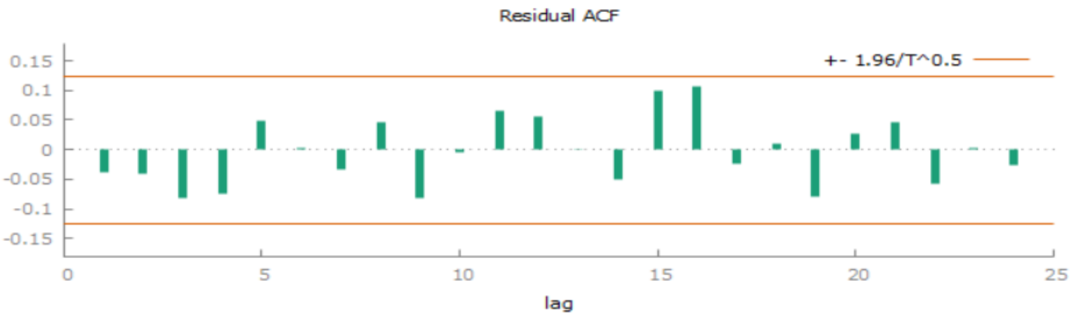
(6) The level-relationship between your variables. Report the estimates of the final ARDL model used for PSS testing. Describe the outcome of main mis-specification tests. Report the outcome of the PSS test. Finally summarise the main facts emerging on this point.

Given that the variables represent quarterly data, we began our analysis by formulating an ARDL(4,4,4,4). However, this model has autocorrelated residuals, meaning that it cannot be used to make estimations. The final model reported below[1] is a model ARDL(4,3,1,0) where we have dropped all the non-significant lags (except for the level variable of unemployment).

Estimates of ARDL(4,3,1,0) model

	coefficient	std. error	t-ratio	p-value	
const	-0.00163401	0.00338002	-0.4834	0.6292	
d_l_dvcft_1	-0.510982	0.0934924	-5.465	1.18e-07	***
d_l_ulc_1	0.289741	0.0834310	3.473	0.0006	***
d_l_dmbs_1	0.109715	0.0219659	4.995	1.15e-06	***
l_qdisi_1	-0.00133373	0.00143207	-0.9313	0.3526	
d_d_l_ulc	0.338197	0.0646114	5.234	3.67e-07	***
d_d_l_ulc_2	0.0826041	0.0358574	2.304	0.0221	**
d_d_l_ulc_3	0.0973294	0.0339188	2.869	0.0045	***
d_d_l_dmbs_1	-0.0376497	0.0180808	-2.082	0.0384	**
d_d_l_dvcft_1	-0.398860	0.0817541	-4.879	1.97e-06	***
d_d_l_dvcft_2	-0.330509	0.0622247	-5.312	2.52e-07	***
d_d_l_dvcft_3	-0.355265	0.0676106	-5.255	3.33e-07	***

Plot of residual ACF



The results of the main misspecification tests are reported as follows:

- The correlogram of the residuals is almost empty, suggesting the absence of autocorrelation.
- Upon conducting the Godfrey test for autocorrelation, the analysis reveals that the residuals across all estimated orders exhibit no evidence of autocorrelation.
- In the context of assessing heteroskedasticity, the application of White's Test yields a p-value of approximately 0. This statistically significant result compels the rejection of the null hypothesis, thereby confirming the presence of heteroskedasticity within the model. Thus, the employment of robust standard errors is warranted to address this issue.
- Regarding the evaluation of residual normality, the Jarque-Bera test produces a p-value of approximately 0. This outcome effectively rejects the null hypothesis that the residuals are normally distributed. As a result, the distribution of residuals deviates from normality.

[1] The model has been estimated with robust standard errors to account for the presence of heteroskedasticity (see White test).

Moving to the results of PSS testing, the speed of adjustment (SOA) is -0.510982 and has a highly statistically significant t-ratio of -5.465. It indicates that approximately 51.10% of any deviation from the equilibrium is corrected each quarter. This signals a relatively fast adjustment back towards the long-term equilibrium after a shock.

The F-test statistic (14.49) exceeds the critical value of 4.35 for I(1) variables at the 5% significance level. Thus, the null hypothesis of absence of a long run relationship is rejected, suggesting that the variables have a long-term relationship. The long run coefficients can be easily computed from the previous outputs:

$$\beta_{\Delta log(ulc)} = \frac{0.289741}{0.510982} \approx 0.567$$

Ceteris paribus, a 1% increase in the unit labour cost is associated with a 0.567% increase in the inflation of value added. This could reflect that higher labour costs translate into higher production costs, which are then passed on to the prices at which output is valued (i.e., the deflator of value-added).

$$\beta_{\Delta log(dmbs)} = \frac{0.109715}{0.510982} \approx 0.215$$

Ceteris paribus, a 1% increase in the deflator of import - which could be seen as a proxy for import prices - is associated with a 0.215% increase in the inflation of value added. The positive relationship implies that increases in the prices of imports (perhaps due to international commodity prices, exchange rate fluctuations, or tariffs) partially feed into domestic inflation as measured by the value-added deflator. This possibly indicates that the domestic economy is somewhat sensitive to import price changes, although the impact is less than proportional.

$$\beta_{log(qdisi)} = -\frac{0.00133373}{0.510982} \approx -0.003$$

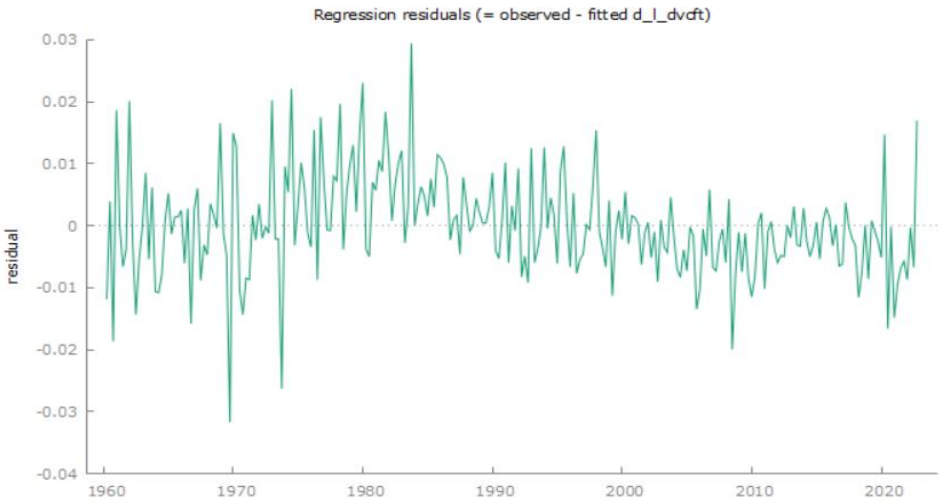
Ceteris paribus, a 1% increase in the unemployment rate is associated with a 0.003% decrease in the inflation of value added. This negative relationship is very small, suggesting that the deflationary effect of unemployment on value-added inflation is negligible and may not be economically meaningful.

(7) Engle-Granger cointegration. Report the estimates of the long run relationship. Plot the residuals of this relationship and report the outcome of their unit root test. Report the estimates of the final ECM model. Summarise main facts here and compare them with ARDL results at point 6

Estimates of long run relationship

	coefficient	std. error	t-ratio	p-value	
const	−0.00377440	0.00467412	−0.8075	0.4201	
d_l_ulc	0.565220	0.0332348	17.01	1.82e-043	***
d_l_dmbs	0.0698283	0.0189051	3.694	0.0003	***
l_qdisi	−0.00353769	0.00189020	−1.872	0.0624	*

Plot of residuals



Outcome of the unit root test of the residuals

```
test without constant
including 3 lags of (1-L)uhat
model: (1-L)y = (a-1)*y(-1) + ... + e
estimated value of (a - 1): -0.548906
test statistic: tau_c(4) = -5.15923
asymptotic p-value 0.001564
1st-order autocorrelation coeff. for e: -0.027
lagged differences: F(3, 243) = 7.942 [0.0000]
```

ECM model

	coefficient	std. error	t-ratio	p-value	
const	5.50344e-05	0.000429774	0.1281	0.8982	
d_d_l_ulc	0.346029	0.0303501	11.40	2.60e-024	***
d_d_l_ulc_2	0.0880136	0.0331108	2.658	0.0084	***
d_d_l_ulc_3	0.103583	0.0325491	3.182	0.0017	***
d_d_l_dmbs	−0.0322790	0.0170548	−1.893	0.0596	*
ECM_1	−0.504147	0.0687624	−7.332	3.54e-012	***
d_d_l_dvcft_1	−0.362351	0.0595961	−6.080	4.74e-09	***
d_d_l_dvcft_2	−0.293164	0.0619235	−4.734	3.78e-06	***
d_d_l_dvcft_3	−0.344985	0.0567545	−6.079	4.78e-09	***

In this framework, to meet the requirements of the Engle-Granger procedure, all the variables are considered as I(1).
In the initial stage of the Engle-Granger procedure, the application of the DF test to the residuals obtained from the static regression, complemented by a visual inspection of the corresponding plots, confirmed the stationarity of the residuals. The final ECM model is shown above. Its estimated speed of adjustment is -0.504, which is statistically significant and indicates the presence of a long-term relationship between the variables.

A comparison between the models in points 6 and 7 is shown in the table below. Values for the SOA and most of the variables are very close, with the exception of *d_l_dmbs*, which is 0.215 in the ARDL estimate and 0.069 in the EG estimate. The results on *l_qdisi* further support our statements about the variable in point 6, i.e. the unemployment rate may not be economically meaningful in the long term.

Parameters	**ARDL**	**ECM**
Speed of Adjustment	-0.511	-0.504
$\beta_{\Delta log(ulc)}$	0.567	0.565
$\beta_{\Delta log(dmbs)}$	0.215	0.069
$\beta_{log(qdisi)}$	-0.003	-0.003

Further analysis involving Engle-Granger cointegration test

As previously noted, the Engle-Granger procedure should be carried out only when all the variables are I(1). Further analyses with the EG procedure were conducted in order to clarify the relationships between the variables and corroborate our conclusions[2].

When considering only the I(1) variables (i.e. *d_l_dvfct* and *l_qdisi*) no long-run relationship emerges (there is no cointegration). Instead, accounting also for the variable *d_l_ulc*, we find a significant estimate, which identifies a long-run relationship between the three variables.

However, in principle, omitting an I(0) variable from the procedure should not change the results of the procedure. Thus, this discrepancy leads us to believe that wage inflation is actually integrated with order one, justifying our initial thoughts emerging from the plots and correlograms of the variable, which showed signs of persistence.
While this could imply that the lack of cointegration was only due to the omission of *d_l_ulc* from the procedure, we also found out that cointegration is still present even when removing *l_qdisi* (testing only with *d_l_dvfct* and *d_l_ulc*).

This could indicate that the only long-run relationship in our model is between domestic and wage inflation, while the unemployment rate does not influence the dependent variable in the long run. As a consequence, we could conclude that the estimator of *l_qdisi* is not super-consistent, due to the lack of cointegration with the dependent variable.

[2] Detailed outputs of these procedures are presented in Appendix A.

(8) Johansen cointegration. Discuss the choice of the deterministic variables in your VAR. Report the rank test results. You might also identify the long run relationship (according to the results with PSS and EG). Summarise by comparing the outcomes here with those reported in points 6 and 7.

In this framework, we decided to estimate a VAR(4) model with its constant as the only deterministic component. This choice is justified by the fact that the explanatory variables represent rates (i.e. internal inflation, wage inflation, import inflation, unemployment). The inclusion of a trend does not seem coherent with the nature of the variables as none of these express a tendency to long-run growth. As with the single equation models, we tested for the absence of autocorrelation in the residuals, together with lag selection criteria, which confirmed that a VAR(4) model is appropriate in this case, coherently with the quarterly nature of the data.

VAR lag selection

lags	loglik	p(LR)	AIC	BIC	HQC
1	2606.96456		-21.291889	-21.004395*	-21.176089*
2	2626.58493	0.00100	-21.321687	-20.804196	-21.113246
3	2650.42399	0.00005	-21.386206	-20.638719	-21.085125
4	2684.37152	0.00000	-21.533922*	-20.556440	-21.140202
5	2698.61184	0.02768	-21.519439	-20.311961	-21.033079
6	2704.98591	0.69108	-21.440213	-20.002739	-20.861213
7	2718.00818	0.05340	-21.415705	-19.748236	-20.744064
8	2730.08215	0.08632	-21.383392	-19.485927	-20.619111

Cointegration test (Johansen)

Corrected for sample size (df = 230)

Rank	Trace test	p-value
0	136.73	[0.0000]
1	64.679	[0.0000]
2	10.720	[0.2366]
3	4.3065	[0.0389]

The Johansen cointegration test suggests a matrix rank equal to 2. This implies the presence of 2 long-run relationships, confirming the results obtained with the PSS and EG procedures.

Identifying the long run relationship

The presence of two long-term co-integration relationships could represent different scenarios, depending on whether they result from the stationary nature of the I(0) variables or from the cointegration relationships between integrated variables I(1). From the univariate analyses we know that a rank equal to 2 is given by the presence of two stationary variables (*d_l_dmbs* and *d_l_ulc*), meaning that there are two long-term cointegration relationships in the model. However, the fact that they refer to stationary variables only tells us that cointegration relationships occur between themselves (that is, the two stationary variables are auto-cointegrated).

From the more detailed analysis made with the Engle-Granger procedure (see paragraph 7), we discovered that *d_l_ulc* may actually be non-stationary, since it is long-term cointegrated with *d_l_dcvft*. The same result is also confirmed by a deepening of the Johansen test[3]. Indeed, by conducting the rank test only between *d_l_dvcft* and *l_qdisi*, a model rank equal to 0 is obtained, confirming the lack of the long-term cointegration between the two integrated variables.

[3] Outputs of this procedure are presented in Appendix B.

Even in this case, by conducting the Johansen test only between *d_dvcft* and *d_l_ulc*, a model's rank equal to 1 is reported, confirming the early long-term cointegration relationship conducted with Engle-Granger.

In conclusion, we believe that the rank equal to 2 is explained by the cointegration between *d_dvcft* and *d_l_ulc*, which adds to the stationary nature of *d_dmbs* (which remains auto-cointegrated).
This assertion also supports the fact that the plot of *d_l_ulc* (presented in paragraph 3) showed a certain degree of persistence and justifies the lesser significance of its stationarity reported in the ADF test.

VECM analysis

```
VECM system, lag order 4
Maximum likelihood estimates, observations 1961:2-2022:4 (T = 247)
Cointegration rank = 2
Case 3: Unrestricted constant

beta (cointegrating vectors, standard errors in parentheses)

d_l_dvcft       1.0000       0.00000
              (0.00000)     (0.00000)
d_l_dmbs       0.00000        1.0000
              (0.00000)     (0.00000)
d_l_ulc       -0.94219       -1.2161
             (0.037930)     (0.18801)
l_qdisi     -0.0040335   -0.0095391
            (0.0018432)  (0.0091361)

alpha (adjustment vectors)

d_l_dvcft     0.031199       0.14568
d_l_dmbs       0.76672      -0.44171
d_l_ulc         1.3292      0.077009
l_qdisi        0.25265      -0.11959
```

To carry out an analysis with a VECM, it seems reasonable to put *d_l_dmbs* first as it is stationary, in order to have a coefficient estimate equal to zero. In this way, the beta estimate of *d_l_ulc* and *l_qdisi* are obtained directly. As observed in the outputs above, the model presents slightly different beta values from those found in the previous models (note however that the sign of the coefficients is reverted in the VECM).
The differences between the models can presumably be explained by the conditions that must be met for the analysis of the single equation dynamic modelling (PSS and EG), i.e. the weak exogeneity of the explanatory (forcing) variables for the long run parameters of interest. Our analysis reveals the absence of weak exogeneity for *d_l_ulc*, as evidenced by the statistically significant speed of adjustment[4], challenging the validity of parameters estimated through single-equation approaches. For these reasons, we must follow the Johansen procedure within the VAR framework.

Moreover (as discussed in point 7), we know that in performing the ECM models (including the VECM) only integrated variables should be taken into consideration: indeed, by excluding *d_l_dmbs* (for which stationarity seems to be certain), we obtain the same results by performing the entire procedure of lag selection, rank test and VECM[5].

[4] Outputs of this procedure are presented in Appendix C.

[5] Outputs of this procedure are presented in Appendix D.

(9) Optional. Filter with HP your variables in p. 1. Report the outcome of the two rank cointegration tests for: (a) the permanent components (b) the transitory components.

After filtering the series with the Hodrick-Prescott filter with $\lambda = 1600$ (again, as we are dealing with quarterly data), we run the two rank Johansen cointegration tests. The outcomes are reported as follows:

a) Johansen cointegration test for the permanent components:

```
Johansen test:
Number of equations = 4
Lag order = 6
Estimation period: 1961:4 - 2022:4 (T = 245)
Case 5: Unrestricted trend and constant

Log-likelihood = 10928 (including constant term: 10232.7)

Rank Eigenvalue Trace test  p-value   Lmax test  p-value
   0    0.23236     115.94 [0.0000]     64.788 [0.0000]
   1    0.11890     51.147 [0.0004]     31.013 [0.0041]
   2   0.065803     20.135 [0.0265]     16.677 [0.0560]
   3   0.014015     3.4580 [0.0629]     3.4580 [0.0630]

Corrected for sample size (df = 219)
Rank Trace test p-value
   0     115.94 [0.0000]
   1     51.147 [0.0003]
   2     20.135 [0.0266]
   3     3.4580 [0.0646]
```

The null hypothesis of an intermediate rank 3 is rejected at the 5% level. This result was expected as it is a consequence of the permanent components being not stationary. However, as these series are smoothed by the filtering process, the Johansen cointegration approach may be uninformative.

b) **Johansen cointegration test for the transitory components:**

```
Johansen test:
Number of equations = 4
Lag order = 4
Estimation period: 1961:2 - 2022:4 (T = 247)
Case 3: Unrestricted constant

Log-likelihood = 3523.03 (including constant term: 2822.07)

Rank Eigenvalue Trace test  p-value   Lmax test  p-value
   0    0.33581     292.85 [0.0000]     101.07 [0.0000]
   1    0.25955     191.78 [0.0000]     74.222 [0.0000]
   2    0.22688     117.55 [0.0000]     63.557 [0.0000]
   3    0.19637     53.998 [0.0000]     53.998 [0.0000]

Corrected for sample size (df = 230)
Rank Trace test p-value
   0     292.85 [0.0000]
   1     191.78 [0.0000]
   2     117.55 [0.0000]
   3     53.998 [0.0000]
```

In principle, knowing that the transitory components are stationary, we can expect to find a maximum rank in the cointegration test. Indeed, all the null hypotheses for intermediate ranks are rejected, meaning that the matrix rank is full.

APPENDIX A - Outputs of Engle-Granger procedures

1) Engle-Granger between d_l_dvcft and l_qdisi

```
Step 1: cointegrating regression

Cointegrating regression -
OLS, using observations 1960:2-2022:4 (T = 251)
Dependent variable: d_l_dvcft

             coefficient   std. error   t-ratio   p-value
  --------------------------------------------------------
  const      -0.0296507    0.00710422   -4.174    4.15e-05 ***
  l_qdisi    -0.0171878    0.00278284   -6.176    2.65e-09 ***

Mean dependent var   0.013925   S.D. dependent var   0.014139
Sum squared resid    0.043341   S.E. of regression   0.013193
R-squared            0.132849   Adjusted R-squared   0.129367
Log-likelihood       731.1917   Akaike criterion    -1458.383
Schwarz criterion   -1451.332   Hannan-Quinn        -1455.546
rho                  0.697859   Durbin-Watson        0.597106

Step 2: testing for a unit root in uhat

Augmented Dickey-Fuller test for uhat
testing down from 4 lags, criterion AIC
sample size 246
unit-root null hypothesis: a = 1

  test without constant
  including 4 lags of (1-L)uhat
  model: (1-L)y = (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.108749
  test statistic: tau_c(2) = -2.30522
  asymptotic p-value 0.3704
  1st-order autocorrelation coeff. for e: -0.004
  lagged differences: F(4, 241) = 17.360 [0.0000]
```

2) Engle-Granger among d_l_dvcft, l_qdisi and d_l_ulc

```
Step 1: cointegrating regression

Cointegrating regression -
OLS, using observations 1960:2-2022:4 (T = 251)
Dependent variable: d_l_dvcft

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const      -0.00460265   0.00478626   -0.9616   0.3372
  l_qdisi    -0.00400109   0.00193351   -2.069    0.0396    **
  d_l_ulc     0.603491     0.0323733    18.64     4.42e-049 ***

Mean dependent var   0.013925   S.D. dependent var   0.014139
Sum squared resid    0.018049   S.E. of regression   0.008531
R-squared            0.638876   Adjusted R-squared   0.635964
Log-likelihood       841.1287   Akaike criterion    -1676.257
Schwarz criterion   -1665.681   Hannan-Quinn        -1672.001
rho                  0.119804   Durbin-Watson        1.745282

Step 2: testing for a unit root in uhat

Augmented Dickey-Fuller test for uhat
testing down from 4 lags, criterion AIC
sample size 246
unit-root null hypothesis: a = 1

  test without constant
  including 4 lags of (1-L)uhat
  model: (1-L)y = (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.433212
  test statistic: tau_c(3) = -4.11996
  asymptotic p-value 0.0172
  1st-order autocorrelation coeff. for e: -0.015
  lagged differences: F(4, 241) = 8.882 [0.0000]
```

3) Engle-Granger between d_l_dvcft and d_l_ulc

```
Step 1: cointegrating regression

Cointegrating regression -
OLS, using observations 1960:2-2022:4 (T = 251)
Dependent variable: d_l_dvcft

             coefficient   std. error    t-ratio    p-value
  ----------------------------------------------------------
  const      0.00520069    0.000686496    7.576    7.00e-013 ***
  d_l_ulc    0.628000      0.0303268     20.71     4.53e-056 ***

Mean dependent var   0.013925   S.D. dependent var   0.014139
Sum squared resid    0.018361   S.E. of regression   0.008587
R-squared            0.632641   Adjusted R-squared   0.631166
Log-likelihood       838.9802   Akaike criterion    -1673.960
Schwarz criterion   -1666.910   Hannan-Quinn        -1671.123
rho                  0.098146   Durbin-Watson        1.790838

Step 2: testing for a unit root in uhat

Augmented Dickey-Fuller test for uhat
testing down from 4 lags, criterion AIC
sample size 247
unit-root null hypothesis: a = 1

  test without constant
  including 3 lags of (1-L)uhat
  model: (1-L)y = (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.511088
  test statistic: tau_c(2) = -4.9267
  asymptotic p-value 0.0002029
  1st-order autocorrelation coeff. for e: -0.030
  lagged differences: F(3, 243) = 10.109 [0.0000]
```

APPENDIX B - Outputs of Johansen cointegration tests

1) Johansen cointegration test between d_l_dvcft and l_qdisi

```
Johansen test:
Number of equations = 2
Lag order = 5
Estimation period: 1961:3 - 2022:4 (T = 246)
Case 2: Restricted constant

Log-likelihood = 2006.45 (including constant term: 1308.34)

Rank Eigenvalue Trace test  p-value  Lmax test  p-value
   0   0.026512      12.111 [0.4475]     6.6099 [0.7187]
   1   0.022115      5.5013 [0.2413]     5.5013 [0.2409]

Corrected for sample size (df = 235)
Rank Trace test p-value
   0     12.111 [0.4518]
   1     5.5013 [0.2407]
```

2) Johansen cointegration test among d_l_dvcft, l_qdisi e d_l_ulc

```
Johansen test:
Number of equations = 3
Lag order = 5
Estimation period: 1961:3 - 2022:4  (T = 246)
Case 3: Unrestricted constant

Log-likelihood = 2811.64 (including constant term: 2113.52)

Rank Eigenvalue Trace test  p-value  Lmax test  p-value
   0     0.22115     73.217 [0.0000]     61.486 [0.0000]
   1    0.026220     11.731 [0.1723]     6.5363 [0.5531]
   2    0.020896     5.1949 [0.0227]     5.1949 [0.0227]

Corrected for sample size (df = 230)
Rank Trace test p-value
   0     73.217 [0.0000]
   1     11.731 [0.1752]
   2     5.1949 [0.0233]
```

3) Johansen cointegration test between d_l_dvcft e d_l_ulc

```
Johansen test:
Number of equations = 2
Lag order = 4
Estimation period: 1961:2 - 2022:4 (T = 247)
Case 2: Restricted constant

Log-likelihood = 2325 (including constant term: 1624.05)

Rank Eigenvalue Trace test  p-value  Lmax test  p-value
   0    0.21804      66.314 [0.0000]     60.751 [0.0000]
   1   0.022268      5.5623 [0.2355]     5.5623 [0.2351]

Corrected for sample size (df = 238)
Rank Trace test p-value
   0     66.314 [0.0000]
   1     5.5623 [0.2349]
```

APPENDIX C - VECM with all variables

VECM outputs of d_l_ulc and l_qdisi

Equation 3: d_d_l_ulc

	coefficient	std. error	t-ratio	p-value	
const	-0.0158996	0.00201254	-7.900	1.13e-013	***
d_d_l_dvcft_1	-0.806588	0.167819	-4.806	2.77e-06	***
d_d_l_dvcft_2	-0.667029	0.142418	-4.684	4.81e-06	***
d_d_l_dvcft_3	-0.249802	0.109282	-2.286	0.0232	**
d_d_l_dmbs_1	-0.0148092	0.0435254	-0.3402	0.7340	
d_d_l_dmbs_2	0.0117169	0.0395626	0.2962	0.7674	
d_d_l_dmbs_3	0.0652049	0.0340563	1.915	0.0568	*
d_d_l_ulc_1	0.245532	0.130872	1.876	0.0619	*
d_d_l_ulc_2	0.181408	0.103793	1.748	0.0818	*
d_d_l_ulc_3	0.0495649	0.0699850	0.7082	0.4795	
d_l_qdisi_1	0.00747064	0.0196779	0.3796	0.7046	
d_l_qdisi_2	-0.0629577	0.0191890	-3.281	0.0012	***
d_l_qdisi_3	-0.0326677	0.0199945	-1.634	0.1037	
EC1	1.32916	0.178183	7.460	1.75e-012	***
EC2	0.0770089	0.0410419	1.876	0.0619	*

Mean dependent var	0.000227	S.D. dependent var	0.016311
Sum squared resid	0.029954	S.E. of regression	0.011387
R-squared	0.542306	Adjusted R-squared	0.512585
rho	-0.011327	Durbin-Watson	2.010814

Equation 4: d_l_qdisi

	coefficient	std. error	t-ratio	p-value	
const	0.000447359	0.00652312	0.06858	0.9454	
d_d_l_dvcft_1	0.591163	0.543942	1.087	0.2783	
d_d_l_dvcft_2	0.627424	0.461611	1.359	0.1754	
d_d_l_dvcft_3	0.162994	0.354207	0.4602	0.6458	
d_d_l_dmbs_1	-0.00306202	0.141076	-0.02170	0.9827	
d_d_l_dmbs_2	-0.0920233	0.128232	-0.7176	0.4737	
d_d_l_dmbs_3	-0.0762223	0.110384	-0.6905	0.4906	
d_d_l_ulc_1	-0.394525	0.424186	-0.9301	0.3533	
d_d_l_ulc_2	-0.333236	0.336418	-0.9905	0.3229	
d_d_l_ulc_3	-0.356702	0.226838	-1.572	0.1172	
d_l_qdisi_1	-0.0274493	0.0637807	-0.4304	0.6673	
d_l_qdisi_2	0.207431	0.0621959	3.335	0.0010	***
d_l_qdisi_3	0.231638	0.0648068	3.574	0.0004	***
EC1	0.252655	0.577532	0.4375	0.6622	
EC2	-0.119595	0.133026	-0.8990	0.3696	

Mean dependent var	0.001543	S.D. dependent var	0.038690
Sum squared resid	0.314688	S.E. of regression	0.036909
R-squared	0.145447	Adjusted R-squared	0.089957
rho	0.038552	Durbin-Watson	1.907832

APPENDIX D - Modelling the VECM

1) VAR lag selection (without d_l_dmbs)

```
VAR system, maximum lag order 8

The asterisks below indicate the best (that is, minimized) values
of the respective information criteria, AIC = Akaike criterion,
BIC = Schwarz Bayesian criterion and HQC = Hannan-Quinn criterion.

lags        loglik    p(LR)       AIC          BIC          HQC

   1    2021.73719           -16.541047   -16.368550*  -16.471567
   2    2038.13019  0.00015  -16.601895   -16.300025   -16.480304
   3    2058.20787  0.00001  -16.693069   -16.261827   -16.519369
   4    2080.69139  0.00000  -16.804044   -16.243430   -16.578234*
   5    2091.43886  0.01063  -16.818427*  -16.128439   -16.540507
   6    2095.58714  0.50456  -16.778495   -15.959135   -16.448465
   7    2102.82840  0.10617  -16.764020   -15.815287   -16.381879
   8    2112.79489  0.01833  -16.771974   -15.693869   -16.337724
```

2) Johansen cointegration test (without d_l_dmbs)

```
Johansen test:
Number of equations = 3
Lag order = 5
Estimation period: 1961:3 - 2022:4 (T = 246)
Case 3: Unrestricted constant

Log-likelihood = 2811.64 (including constant term: 2113.52)

Rank Eigenvalue Trace test  p-value  Lmax test  p-value
   0    0.22115     73.217 [0.0000]     61.486 [0.0000]
   1   0.026220     11.731 [0.1723]     6.5363 [0.5531]
   2   0.020896     5.1949 [0.0227]     5.1949 [0.0227]

Corrected for sample size (df = 230)
Rank Trace test p-value
   0     73.217 [0.0000]
   1     11.731 [0.1752]
   2     5.1949 [0.0233]
```

3) VECM (without d_l_dmbs)

```
VECM system, lag order 5
Maximum likelihood estimates, observations 1961:3-2022:4 (T = 246)
Cointegration rank = 1
Case 3: Unrestricted constant

beta (cointegrating vectors, standard errors in parentheses)

d_l_dvcft       1.0000
              (0.00000)
d_l_ulc       -0.94684
             (0.036716)
l_qdisi     -0.0042745
            (0.0017686)

alpha (adjustment vectors)

d_l_dvcft      0.19011
d_l_ulc         1.4956
l_qdisi        0.27628
```