3. DATA OF THE MODEL

3.1 Data Sources

The estimated model is supposed to represent the Spanish economy. The structure of the model closely follows the structure of national accounts, which is a consistent system of time series for the whole economy and its sectors. Thus, the national accounts serve as the best data base for the model. As the key focus of the model is on short term forecasts and cyclical effects of shocks and economic policy measures, *quarterly* national accounts (QNA) are chosen here. Since the introduction of the European System of Accounts 1995 (ESA 1995) the Spanish statistical institute (INE) has regularly published raw, seasonally and calendar adjusted quarterly national accounts series as well as their trend-cycle component[1]. The INE obtains quarterly raw national accounts data by applying the Chow-Lin method (Chow and Lin 1971) to the annual aggregates. This is a procedure for temporal disaggregation with the help of suitable indicator series. Details of the INE's approach are described in Quilis (2001). The INE subsequently adjusts the series applying TRAMO and SEATS (Gómez and Maravall 1996). However, in the model described here raw data are used. Whereas most other quarterly models rely on seasonally adjusted data, there are good reasons for using raw data in the present case:

- The model was originally designed as part of a multi-country-model of the euro area. The use of seasonally unadjusted data allows for direct seasonal adjustment of the euro area aggregates, which is preferable to indirect seasonal adjustment, if the focus is on the euro area[2].
- The series contain more information. Structural breaks in the data can be detected more easily and are not watered down by seasonal adjustment.

---

[1] Under the ESA79 quarterly data were limited to the trend-cycle component.

[2] For details see Rietzler, Stephan, Wolters (2000, 2001)

The use of the model for repeated forecasts and policy simulations implies that the data base should be easily accessible and easily updated. For this purpose the quarterly national accounts data, which are published roughly 50 days after the end of the latest quarter are equally ideal. Although Spain's quarterly accounts data offer quite a number of time series, they are not sufficient for a full macro model of the Spanish economy. A sectoral breakdown of data is only available from *annual* national accounts. Often these series begin only in 1995, which means that for the estimation period of the model (1980-2002) only eight years of data exist. For its Economic Outlook the OECD has complemented the annual national accounts basis of the INE with its own estimates of data, which is not available from the INE. In particular the OECD has provided estimates for the period prior to 1995 for a number of annual time series including for example disposable income of households. In addition the Spanish quarterly national accounts still do not offer a regional breakdown of foreign trade.

Consequently some quarterly time series had to be constructed by temporal disaggregation and other calculations. In order to keep the data base within manageable proportions, such methods were only applied in the most urgent cases. On the one hand this approach entails some serious limitations for the model, but on the other hand, it helps to keep the procedures transparent and to make an update of the data base less time-consuming.

As the highest frequency of the OECD Economic Outlook is half-yearly data (seasonally adjusted), the author has carried out her own temporal disaggregation of some annual time series taken from the OECD Economic Outlook No. 74:

- Disposable income of households via subseries.
- Total demand (i.e. $GDP + imports = consumption + investment + exports$) of small euro area countries with insufficient quarterly data
- Government investment.

Indicator series for the temporal disaggregation have usually been taken from the QNA. In some cases like the temporal disaggregation of GDP and imports of Greece, Portugal and Ireland indicator series from the OECD's Main Economic Indicators were used. When no suitable indicator existed, the method without indicator was applied. Details of the procedure are described in the Appendix A.3.

An alternative possibility would have been to try and obtain the respective series and some additional ones (such as the tax wedge and the replacement ratio) from the Bank of Spain, whose experts have constructed numerous time series for their own model. However, there is no extensive documentation of how these series were derived nor is there any guarantee that they will be available in the future. An additional drawback is that they are all seasonally adjusted.

For the foreign trade block real effective exchange rates had to be calculated and the exports of goods had to be broken down by regions. For these calculations, which are described in detail in Appendices A.4 and A.5, consumer price indices of various countries were taken from the OECD Main Economic Indicators. Country weights were determined with the help of annual data from the IMF's directions of trade statistics. Monthly trade data from the Ministry of the Economy[3] were used for breaking down Spain's exports of goods according to the following destinations: EMU, EU-15 outside EMU (i.e. United Kingdom, Sweden, Denmark), USA, the rest of the world.

Additional time series, which are mostly exogenous to the model, have been taken from the IMF and the OECD. Among others they include the oil price (per barrel Brent), nominal exchange rates and interest rates.

The scope of the model is limited by the availability of suitable data. This is particularly true for the government sector, for which - with only a few exceptions - *annual* time series begin only in 1995[4]. Quarterly data for the whole estimation period are only available for government consumption, taxes and subsidies on products as well as taxes less subsidies on production and imports. An annual series of government investment at constant prices is published with the OECD Economic Outlook and can be temporally disaggregated. As a consequence a large part of government demand is included in the model, but it is not possible to make statements about the fiscal deficit or the stock of government debt.

Generally the inclusion of stocks in the model proved difficult. Statistical sources for series such as the capital stock and wealth are scarce and not easily compatible

---

[3] aggregated into quarters

[4] Extended annual government time series - partly beginning in 1964 have become available from the OECD after the completion of the model estimations.

with the national accounts framework. The main economic impact of stocks comes with their rapid revaluation, e.g. in a stock market crash or a house price bubble. These revaluations strongly affect investment and consumption decisions and it would be desirable to have suitable equations for them. However, it is well known, that e.g. stock prices follow a random walk and cannot be modelled well or even be forecast. Thus, any model including stocks would miss a large part of their effects.

For a full list of variables used in the model c.f. Appendix A.1.1. The list provides the abbreviations used in the model equations, explanations of the series as well as their sources. All series are available at least for the period since 1980. In addition, the time series that appear in the estimated behavioural equations are shown in graphs in the same appendix. Their stochastic properties are given in section 3.3 below. As the error-correction equations are usually estimated with logarithms, graphs and unit root test results generally refer to logs. The latter are denoted by the variable abbreviations in *lower case* letters.

3.2 Structural Breaks in the Data

In Chapter 2 the main economic developments and institutional changes have been summarised. These alone render structural breaks in the time series hardly surprising. The succession of so many different monetary regimes alone would justify quite a lot of instability in the data.

However, reforms and changing monetary regimes obviously cannot explain the structural breaks in the series. Rather, there seem to be purely statistical reasons, which manifest themselves in

- Level shifts
- Changes of the seasonal patterns
- Breaks in the trend

The structural breaks enumerated above occur in a number of series - often in combination. Graphs of the logarithms of private final consumption expenditure at constant prices of 1995 (cp95), the private consumption deflator (pc) and unit labour cost (ulc) shall illustrate these structural breaks:

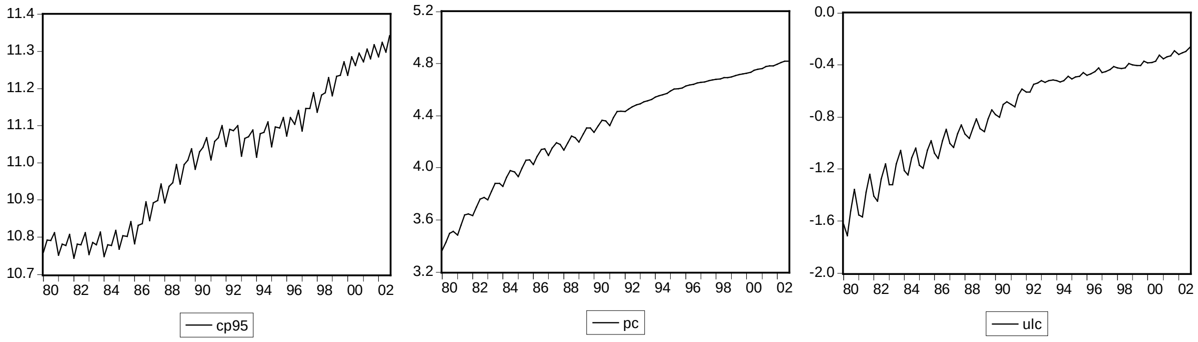

Figure 3.1: Examples of structural breaks in the time series

Private consumption expenditure shows a downward level shift at the beginning of 1993 as well as a change of the seasonal pattern at the end of the 1990s. The private consumption deflator has a very pronounced seasonal pattern until the end of 1991. Unit labour cost exhibits a combination of a change of the seasonal pattern and a change of the trend slope.

Unlike the structural breaks in the German series due to reunification, the level shifts and changes of seasonal patterns in the Spanish series do not seem to be the consequences of shifts in the real economy. More probably they point to statistical problems in the compilation of the national accounts data.

In the estimations the structural breaks can be modelled with dummy variables. Changes of the seasonal patterns (as long as they occur once and abruptly) and level shifts thus do note pose serious problems for the estimations.

Breaks in the trend are not problematic, if they also appear in the explanatory variables of an equation and thus do not have to be modelled with a broken deterministic trend. The use of deterministic trends in the estimations entails the risk of large over- or underestimations in the case of a renewed trend shift. In addition to the structural breaks mentioned above, there are also numerous outliers.

3.3 Results of the Unit Root Tests

Before any equation is estimated, it has be ensured that it is balanced. This means that we have to have the same order of integration on each side of the equation.

Therefore all series have to be tested for unit roots. Depending on the deterministics of the series the author applies either the augmented Dickey-Fuller Test (Dickey and Fuller 1979) or the Perron Test (Perron 1989, Perron and Vogelsang 1993). The latter has been devised to avoid spurious results due to breaks in the series (for details about the unit root tests see Appendix A.2).

These tests give a first indication of the stochastic properties of the series. They cannot be interpreted as absolute truth. There may be contradictory results, for example, when cointegration is found between series of different order of integration, as it is the case for the GDP deflator and nominal wages. According to the unit root tests these two series cannot be cointegrated, because the GDP deflator is classified as I(2), whereas the nominal compensation per employee is assumed to be stationary with respect to a broken deterministic trend. However, cointegration is found, when the structural break is modelled explicitly.

For these reasons the author follows a pragmatic approach, where the unit root tests, theoretical considerations and the results of cointegration tests are all taken into account.

It also has to be kept in mind that the results of the unit root tests hold only for the period tested. If the reference period is changed, the integration properties of the series may also change.

Tables 3.1 and 3.2 show the results of the unit root tests carried out. These include only those variables that appear in estimation equations. Variables which are only relevant for definitions are not tested, because their stochastic properties do not matter for modelling. As all series end in 2002Q4 only the beginning of the sample is given in the tables. In the table c stands for constant, t for trend and s for centred seasonal dummies. TS is used as an abbreviation for trend stationary variables. Variable names in lower case letters refer to the logs of the variables. The 5-% critical values are given in brackets.

Table 3.1: Results of the ADF-Tests

Variable	From	Deterministics	Dynamics	Test statistics	Result
cgov95	1981Q2	ct, s*SD9201i	lags 3,4	-1.71 (-3.46)	I(1)
Δ cgov95	1981Q2	c, s*SD9201i	lag 3	-10.36 (-2.90)	
coeosmin	1981Q2	ct, s	lag 4	-2.37 (-3.46)	I(1)
Δ coeosmin	1981Q4	c, s	lags 1,4,5	-4.96 (-2.90)	
coeosmin	*1986Q1*	ct, s	lag 4	-2.38 (-3.46)	I(1)
Δ coeosmin	*1986Q1*	c, s	lags 1,4,5	-3.87 (-2.90)	
cp95	1981Q2	ct, s	lags 1-4	-3.25 (-3.46)	I(1)/I(2)
Δ cp95	1981Q2	c, s	lags 1-3	-2.27 (-2.90)	
cp95	*1986Q1*	ct, s	lag 4	-3.26 (-3.48)	I(1)
Δ cp95	*1986Q1*	c, s	lags 2,3,4	-6.20 (-2.90)	
cpi	1981Q2	ct, s	lag 4	-3.60 (-3.46)	I(2)
Δ cpiewu	1981Q2	c, s	lags 1-3	-2.20 (-2.90)	
cpiewu	1981Q2	ct	lag 4	-5.34 (-3.46)	I(2)
Δ cpiewu	1981Q2	c	lags 1-3	-2.26 (-2.90)	
(cpiewu+ecu)	1980Q3	ct	lag 1	-2.87 (-3.46)	I(1)
Δ (cpiewu+ecu)	1980Q3	c	lag 0	-6.74 (-2.90)	
ecu	1980Q3	c	lag 1	-1.96 (-2.89)	I(1)
Δ ecu	1980Q3	none	lag 0	-6.24 (-1.94)	
ee	1981Q2	ct, s	lags 1,2,4	-3.44 (-3.46)	I(1)
Δ ee	1980Q4	c, s	lag 1	-3.00 (-2.89)	
ee	*1986Q1*	ct, s	lags 1,2,4	-3.33 (-3.46)	I(1)/I(2)
Δ ee	*1986Q1*	c, s	lag 1	-2.13 (-2.90)	
es	1982Q2	ct	lags 4,8	-3.03 (-3.46)	I(1)
Δ es	1982Q2	c	lags 1-5,7	-6.03 (-2.90)	
ewuoes_dtot	1981Q2	ct, s	lag 4	-2.69 (-3.46)	I(1)
Δ ewuoes_dtot	1981Q2	c, s	lag 3	-9.09 (-2.90)	

Variable	From	Deterministics	Dynamics	Test statistics	Result
gdp95	1981Q2	ct, s	lags 1,4	-2.31 (-3.46)	I(1)
Δ gdp95	1981Q2	c, s	lags 1-3	-3.29 (-2.90)	
gdp95	*1986Q1*	ct, s	lags 1,4,5	-2.78 (-3.46)	I(1)
Δ gdp95	*1986Q1*	c,s	lags 2,3	-11.92 (-2.90)	
icon95	1981Q2	ct,s	lags 1,4	-2.83 (-3.46)	I(1)/I(2)
Δ icon95	1981Q2	c, s	lags 1-3	-2.32 (-2.90)	
icon95	*1986Q1*	ct,s	lag 4	-2.49 (-3.46)	I(1)/I(2)
Δ icon95	*1986Q1*	c, s	lags 1-3	-2.20 (-2.90)	
ifc	1981Q3	ct, s	lags 1,3,4	-2.26 (-3.46)	I(1)
Δ ifc	1981Q1	c, s	lags 1,2	-3.35 (-2.90)	
ifc	*1986Q1*	ct, s	lags 1,4	-2.22 (-3.46)	I(1)
Δ ifc	*1986Q1*	c, s	lag 0	-5.95 (-2.90)	
imeq95	1981Q2	ct, s	lags 2-4	-3.01 (-3.46)	I(1)
Δ imeq95	1981Q3	c, s	lags 1,4	-4.60 (-2.90)	
imeq95	*1986Q1*	ct, s	lags 2,4	-2.30 (-3.46)	I(1)
Δ imeq95	*1986Q1*	c, s	lag 1	-3.92 (-2.90)	
IS95	1981Q1	c, s	lag 3	-10.92 (-2.90)	I(0)
m95	1981Q2	ct, s	lag 4	-2.38 (-3.46)	I(1)
Δ m95	1981Q2	c, s	lags 1-3	-3.76 (-2.90)	
NL	1982Q1	c	lags 1,7	-1.49 (-2.90)	I(1)
Δ NL	1982Q1	none	lags 1,2,6	-5.79 (-1.94)	
NL	*1986Q1*	c	lags 1,7	-1.07 (-2.90)	I(1)
Δ NL	*1986Q1*	none	lags 3,6	-5.15 (-1.94)	
NS	1981Q4	c	lags 1, 2 6	-0.90 (-2.90)	I(1)
Δ NS	1980Q4	none	lag 1	-6.39 (-1.94)	
oil$	1981Q2	c	lags 1-3	-2.99 (-2.90)	I(0)/I(1)
Δ oil$	1981Q2	none	lags 1-3	-8.56 (-1.94)	
pc	1981Q2	ct, s*SD9201i	lags 1,4	-3.47 (-3.46)	I(2)
Δ pc	1981Q2	c, s*SD9201i	lags 1-3	-1.53 (-2.90)	

Variable	From	Deterministics	Dynamics	Test statistics	Result
pgdp	1981Q2	ct, s	lag 4	-3.02 (-3.46)	I(2)
Δ pgdp	1981Q2	c, s	lags 1-3	-1.44 (-2.90)	
pgdppm	1981Q2	c	lag 4	-1.55 (-2.90)	I(1)
Δ pgdppm	1981Q3	none	lag 4	-7.37 (-1.95)	
pifc	1981Q3	ct, s	lags 4,5	-3.86 (-3.46)	I(1)/I(2)
Δ pifc	1981Q3	c, s	lags 1-4	-2.74 (-2.90)	
pm	1981Q2	ct	lag 4	-2.89 (-3.46)	I(1)
Δ pm	1981Q2	c	lags 1-3	-3.51 (-2.90)	
prodet	1981Q2	c, s	lags 1-4	-2.77 (-2.90)	I(1)
Δ prodet	1981Q3	s	lags 1-4	-2.18 (-1.94)	
px	1981Q2	ct, s	lags 2,4	-3.74 (-3.46)	I(2)
Δ px	1981Q2	c, s	lags 1-3	-2.17 (-2.90)	
raw	1980Q3	c	lag 1	-1.68 (-2.89)	I(1)
Δ raw	1981Q3	none	lag 4	-6.22 (-1.94)	
raw	*1986Q1*	c	lags 1,5	-2.96 (-2.90)	I(0)/I(1)
Δ raw	*1986Q1*	none	lag 4	-6.22 (-1.94)	
rawewu	1980Q3	c	lag 1	-1.96 (-2.89)	I(1)
Δ rawewu	1980Q3	none	lag 0	-8.03 (-1.94)	
rawreu	1981Q3	c	lags 1,5	-1.48 (-2.90)	I(1)
Δ rawreu	1981Q3	none	lag 4	-7.55 (-1.94)	
rawus	1980Q3	c	lag 1	-1.74 (-2.89)	I(1)
Δ rawus	1980Q3	none	lag 0	-6.86 (-1.94)	
reu_dtot95	1981Q2	ct, s	lag 4	-1.83 (-3.46)	I(1)
Δ reu_dtot95	1980Q3	c, s	lag 0	-8.72 (-2.89)	
rweepgdp	1981Q2	ct, s, s*SD9201i	lag 4	-3.55 (-3.46)	TS
Δ rweepgdp	1981Q4	c, s, s*SD9201i	lags 1-5	-4.23 (-2.90)	
tind	1981Q1	ct, s, s*SD9101	lags 1-3	-3.14 (-3.46)	I(1)
Δ tind	1981Q1	c, s, s*SD9101	lags 1,2	-12.21 (-2.90)	

Variable	From	Deterministics	Dynamics	Test statistics	Result
tind	*1986Q1*	ct, s, s*SD9101	lag 3	-8.28 (-3.46)	TS
Δ tind	*1986Q1*	c, s, s*SD9101	lags 1,2	-11.50 (-2.90)	
u	1981Q1	c, s	lags 1,3	-1.98 (-2.90)	I(1)
Δ u	1981Q3	s	lags 1,2,4	-2.91 (-1.94)	
UR	1981Q1	c, s	lags 1,3	-2.05 (-2.90)	I(1)
Δ UR	1981Q1	s	lags 1,2	-2.53 (-1.94)	
UR	*1986Q1*	c, s	lags 1,3	-1.55 (-2.90)	I(1)
Δ UR	*1986Q1*	s	lags 1,2	-2.21 (-1.94)	
UR1	1981Q1	c, s	lags 1,3	-2.05 (-2.90)	I(1)
Δ UR1	1981Q1	s	lags 1,2	-2.53 (-1.94)	
us_dtot95	1980Q3	ct	lag 1	-2.38 (-3.46)	I(1)
Δ us_dtot95	1980Q3	c	lag 0	-6.42 (-2.89)	
usd	1980Q3	c	lag 1	-2.52 (-2.89)	I(1)
Δ usd	1980Q3	none	lag 0	-6.57 (-1.94)	
x95	1981Q2	ct, s	lags 1,4	-1.95 (-3.46)	I(1)
Δ x95	1981Q2	c, s	lags 2,3	-15.11 (-2.90)	
xg95	1982Q1	ct, s	lags 1-5,7	-1.97 (-3.46)	I(1)
Δ xg95	1982Q1	c, s	lags 1,3,6	-8.93 (-2.90)	
xg95	*1986Q1*	ct, s	lags 4,6	-4.39 (-3.46)	I(1)
Δ xg95	*1986Q1*	c, s	lags 1,3,6	-8.30 (-2.90)	
xg95ewu	1981Q2	ct, s	lags 1,2,4	-2.05 (-3.46)	I(1)
Δ xg95ewu	1981Q2	c, s	lags 1,3	-11.13 (-2.90)	
xg95row	1980Q3	ct, s	lag 1	-1.31 (-3.46)	I(1)
Δ xg95row	1980Q3	c, s	lag 0	-12.58 (-2.89)	
xg95row	*1986Q1*	ct, s	lags 1,4	-5.29 (-3.46)	TS
Δ xg95row	*1986Q1*	c, s	lags 2,4	-10.60 (-2.90)	
xg95us	1982Q1	ct, s	lags 1,7	-2.00 (-3.46)	I(1)
Δ xg95us	1982Q1	c, s	lag 6	-12.05 (-2.90)	

Variable	From	Deterministics	Dynamics	Test statistics	Result
xgicon95	*1986Q1*	ct, s	lag 4	-3.70 (-3.46)	I(1)
Δ xgicon95	*1986Q1*	c, s	lag 4	-9.88 (-2.90)	
ximeq95	1981Q2	c,t,s	lags 1,4	-2.05 (-3.46)	I(1)
Δ ximeq95	1981Q2	c,s	lags 2,3	-13.02 (-2.90)	
xs95	1982Q2	ct, s	lags 2,4-6,8	-2.10 (-3.46)	I(1)/I(2)
Δ xs95	1982Q2	c, s	lags 1-3,6,7	-1.83 (-2.90)	
xs95	*1986Q1*	ct, s	lag 4	-2.30 (-3.46)	I(1)
Δ xs95	*1986Q1*	c, s	lags 1,4,5	-5.20 (-2.90)	
yd95	1981Q2	ct, s, s*SD	lags 1,4	-2.41 (-3.46)	I(1)/I(2)
Δ yd95	1981Q2	c, s, s*SD	lags 1-3	-2.77 (-2.90)	
yd95	*1986Q1*	ct, s, s*SD	lag 4	-2.10 (-3.48)	I(1)
Δ yd95	*1986Q1*	c, s, s*SD	lag 3	-7.81 (-2.90)	

All series in Tables 3.1 and 3.2 begin in 1980 and thus include 92 observations. The estimation periods of the ADF-Tests are usually shorter due to lags and differencing. In cases, where the estimation period of the model equations begins later (in 1986, the year of EU accession), the ADF test equally has to be carried out for the shorter period.

Most series, which were tested, are integrated of order one i.e. I(1), which means that they are stationary in differences. The only series, which is unambiguously stationary in levels is the change of inventories plus net acquisition of valuables, which fluctuates around zero. The oil price is a borderline case.

In some cases the ADF-Test indicates integration of order 2. Price indices like the GDP deflator, the private consumption deflator and the export deflator are typical cases. As price indices are often found to be I(2), this result is accepted. In the case of employment or construction investment this result is dismissed as spurious.

The majority of the series tested with the Perron Test (Table 3.2) is found to be trend stationary. In six cases the null hypothesis of a unit root in the series cannot be rejected: consumption of fixed capital (cfc) from 1980-2002, compensation of employees plus operating surplus and mixed income (coeosmin) from 1986-2002, the government consumption deflator (pcgov), direct taxes paid by households (td) from 1986-2002,

Table 3.2: Results of the Perron Tests

Variable	Sample	Model	Test statistics	Result
cfc	1980Q1	Model C, $\lambda = 0.6$	-2.60 (-4.24)	I(1)
cfc	1986Q1	Model A, $\lambda = 0.4$	-5.64 (-3.72)	TS
coeosmin	1980Q1	Model C, $\lambda = 0.6$	-6.47 (-4.24)	TS
coeosmin	1986Q1	Model C, $\lambda = 0.4$	-4.07 (-4.22)	I(1)
pcgov	1980Q1	Model C, $\lambda = 0.6$	-2.18 (-4.24)	I(1)
rwee	1980Q1	Model C, $\lambda = 0.6$	-5.83 (-4.24)	TS
rwee	1986Q1	Model C, $\lambda = 0.4$	-4.73 (-4.22)	TS
rweepgdp	1980Q1	Model C, $\lambda = 0.6$	-7.14 (-4.24)	TS
sc	1980Q1	Model B, $\lambda = 0.6$	-5.02 (-3.94)	TS
spread	1980Q1	Model A, $\lambda = 0.7$	-5.13 (-3.80)	TS
td	1980Q1	Model B, $\lambda = 0.6$	-6.39 (-3.94)	TS
td	1986Q1	Model B, $\lambda = 0.4$	-1.59 (-3.91)	I(1)
trr	1980Q1	Model B, $\lambda = 0.6$	-3.56 (-3.94)	I(1)
ulc	1980Q1	Model C, $\lambda = 0.6$	-5.08 (-4.24)	TS
wee	1980Q1	Model C, $\lambda = 0.6$	-4.40 (-4.24)	TS
xg95reu	1980Q1	Model B, $\lambda = 0.6$	-2.98 (-3.94)	I(1)

transfers received by households (trr) and exports to the EU-15 outside the euro area at prices of 1995 (xg95reu).